Faculdade de Engenharia da Universidade do Porto

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Modeling a Decentralized Market-Based Scheme for Responsive Demands

Tiago de Sousa Garcia

Dissertação realizada no âmbito do
Mestrado Integrado em Engenharia Eletrotécnica e de Computadores
Major Energia

Orientador: Prof. Doutor João Paulo da Silva Catalão
Co-orientador: Prof. Doutor Miadreza Shafiekhah

Julho 2017

Abstract

Nowadays, most electricity users bid into the market putting an infinite value on energy in which concerns to the price to pay for it, without taking profit of the possibility to change their consumption to hours in which the cost of energy is lower. Demand response programs allow users to interact with the market and make changes in their consumption profiles as to diminish the cost of purchasing energy. Previous developed demand response models based on a bidding process, such as the centralized one, entail problems such as scalability or privacy, and decentralized models relying on unidirectional communication entail problems such as avalanche effects. To overcome those problems, a new decentralized model relying on bidirectional communication was developed in this work. In this model, each user bids into the market according to their consumption urgency and a set of parameters defined by a reinforcement learning algorithm. The bids are aggregated at the intermediate level by an aggregator agent, who then communicates the aggregate bids into the market. After the market clearing price is determined, it is communicated to the aggregator and then to the end-users, who buy energy according to the market price and their previously defined demand bids. The goal is to minimize the costs of buying energy. Three demand response level were considered as to analyze the impact of the demand response level in the cost of purchasing energy. The obtained results are compared to the results obtained with a centralized model and an uncontrolled approach. Results demonstrate that the decentralized model of this work reduces considerably the costs of purchasing energy when comparing to an uncontrolled approach, and almost as optimally as a centralized model.

Key-words: Multi-Agent System; Bidding Strategy; Decentralized Market-Based Scheme; Demand Response.

Resumo

Nos dias correntes, a grande maioria dos consumidores de energia definem a sua procura de energia atribuindo um valor infinito à energia no que diz respeito ao valor a pagar por esta, não tirando partido da possibilidade de alterar o consumo de energia para outras horas de menor custo energético. Programas de demand response permitem aos utilizadores de energia a sua interação com o mercado de energia, fazendo estes alterações nos seus perfis de consumo de forma a diminuir o custo de compra de energia. Modelos de demand response baseados em licitações, tal como o modelo centralizado, estão associados a problemas tais como a escalabilidade e a privacidade da informação dos consumidores, sendo que modelos descentralizados baseado em comunicação unidirecional estão associados a problemas tais como os efeitos avalanche. De forma a ultrapassar os mencionados problemas, foi desenvolvido neste trabalho um modelo descentralizado baseado em comunicação bidirecional. Neste modelo, cada consumidor faz uma licitação ao mercado energético de acordo com a sua urgência de consumo energético e um conjunto de parâmetros definidos através de um algoritmo de reinforcement learning. As licitações são agregadas no nível intermédio por um agente agregador, o qual posteriormente comunica as licitações ao mercado. Após determinação do preço por unidade de energia, este é comunicado ao agente agregador, e posteriormente aos consumidores, os quais compram energia de acordo com este mesmo preço e as suas curvas de procura previamente definidas. O objetivo deste trabalho consiste em minimizar o custo de compra de energia. Três níveis de demand response foram considerados de forma a analisar o impacto do nível de penetração de demand response no custo de compra de energia. Os resultados são comparados com os resultados obtidos através de um modelo centralizado, bem como com os resultados obtidos num cenário de carga totalmente inflexível. Os resultados demonstram que o modelo descentralizado desenvolvido neste trabalho reduz consideravelmente o custo de compra de energia quando comparado com um cenário de carga totalmente inflexível, e de forma quase tão otimizada como um modelo centralizado.

Palavras-chave: Sistema Multi-Agente; Estratégia de Licitação; Esquema Base-Mercado Descentralizado; Demand Response.

Acknowledgments

Regarding all these months of work, I would like to thank Prof. Dr. Miadreza Shafie-khah for all the helpful guidance, advices and availability during this process, Dr. Gerardo Osório for all the feedback, encouragement and availability, and Prof. Dr. João Catalão for the opportunity to work on this theme with such a great team.

I would also like to thank my family, specially my parents, who are probably prouder of my work than I can ever imagine, and all my friends, who never fail to support me even when I tend not to believe in myself. For that I will be eternally grateful.

I also would like to thank all friends I've been making all those years with whom I shared ideas and love that shaped me.

Last but not least, I would like to thank all people who indirectly inspire me and mankind in general to do well.

Porto, June 20, 2017

Tiago Garcia

Contents

List of Acronyms

DR	Demand Response
HVAC	Heating, Ventilation and Air Conditioning
MAS	Multi-Agent System
MBC	Market-Based Control
MCP	Market Clearing Price
OPF	Optimal Power Flow

List of Figures

List of Tables

Chapter 1

Introduction

One of the major problems in the present era consists on carefully and wisely explore the resources that can be obtained throughout interaction and exploration of the present resources on our planet. As for that, it is important not only to explore resources efficiently, but also to explore them in a way that considers consistency on its exploration.

A problem that energy networks are facing nowadays is the high and rising consumption level, as mentioned in [1]. Since not only population is growing exponentially, but also due to the fact that electricity is substituting other resources whose usage was almost electricity independent. Thus, it is possible to observe an increment in the requirements of energy and the introduction of problems to the network control, such as the reduction of the possibility for the network agents to find an alternative during peak hours to fulfill the energy requirements of the fleet in a scenario in which generators fail to work. To solve that problem, it would be crucial to reduce consumption during peak hours, postponing consumption to valley hours, where the network agents won't face problems such as a demand level superior the maximum accepted demand, i.e., maximum amount of energy the market can provide at a certain time. Such problems can be solved using demand response programs.

Demand response programs allow end-use customers to make changes in electric usage from their usual consumption patterns, allowing them to postpone load consumption from peaks hours to valley hours. In [2], DR is defined as a mechanism that enables customers to participate in the electric market as to improve power system efficiency and integrate renewable generation. However, according to [3], few DR programs are developed for residential customers. Nevertheless, the development of technologies related to the smart grids, such as smart metering, were considerably developed during last years, leading the residential DR to increase in terms of attractiveness due to the potential these programs may bring, as mentioned in [4]. As mentioned in [5], a smarter control of the residential loads will lead to a more efficient energy management strategy for the customers.

As an incentive for the customers, the price to pay for energy during the valley hours is lower than the price during peak hours. Allied to that fact, the higher the quantity demanded, the higher the price to pay for an energy unit. Therefore, demand response will give customers the possibility to have a more balanced load pattern throughout the day, reducing the energy usage cost and reducing consumption during peak hours.

Some DR strategies are presented in [6-9] as to solve problems related to peak load shedding problems and shifting problems. Other strategies used in the United States for commercial and industrial customers are direct load control, real-time pricing, and time-of-use programs, as mentioned in [10].

The goal of this thesis is to model a decentralized scheme considering responsive demands, i.e., implementing a market strategy in which the DR-enabled customers will bid into the market according to their energy requirements, and the market will answer according to their bids, similar to the model presented in [11] for electric vehicles.

This strategy will be based on a cost-minimization strategy of the consumer's consumption. Two of those strategies are the centralized and the decentralized approaches, mentioned in [11]. In which concerns to the results, regarding the energy costs, both schemes lead to good results. The central model considers a central agent, denoted as the aggregator, that controls the consumption of the end-users directly, guaranteeing the satisfaction of the demand and the cost minimization. Centralized schemes rely on bi-directional communication between the end-users and the aggregator agent. For them to participate in the market, the DR-enabled customer's information is sent to a central level, where the energy consumption is scheduled and the control signals from the aggregator are sent to the users. Despite the fact that this approach leads to optimal costs, it is not useful for large fleets, since a huge number of users would lead to complex fleets, being the information difficult for them to process, as mentioned in [11]. To overcome the mentioned problem, the fleet constraints will be represented in an aggregated way such as in [12-16].

On the other side, the decentralized approaches have lower communication requirements, being enough to broadcast the price in order to control the consumption, as demonstrated in [17-20] in the approach used for electric vehicles. However, undesirable outcomes can occur, like the possibility of simultaneous reactions, such as the avalanche effects, or errors in forecasting the consumers' attitude in relation to the price signals. This happens due to the fact that the impact of the DR-enabled customer's bids is not taken into account. To avoid the mentioned undesirable outcomes, bidirectional communication will be considered in the decentralized approach. Despite the rising of the communication requirements, the load synchronization problems will be overcome by the usage of a bidding process, such as in [21], or an iterative process such as in [22-23]. Each end-user places its bids according to their energy requirements, and the market answers to the bids by communicating the clearing price.

The customer's sensitivity to the market prices will be represented in this work by the introduction of a bidding process similar to the one presented in [11], in which each end-user will define its energy demand bid according to its requirements. It will be established a multi-agent system, in which each consumer will optimize its cost, similarly to the decentralized model. The system will be modeled within three levels. In the lower level, each customer will represent its willingness to buy energy by placing bids that represent the status of the consumers.

At the intermediate level, bids will be aggregated by the aggregator agent. In the upper level, aggregated bids are received and the market clearing will be performed. The market clearing price will be communicated to the aggregator agent, that will inform it the customers, leading them to buy energy according to their demand bid curves, being those bids represented in a three-step piecewise linear function such as in [24].

The different agents' bids are divided in three steps. The first one represents the inflexible demand, i.e., the critical loads as defined in [25], being the agents able to pay any amount of money to serve those loads. The second and third steps exist to allow the agents to buy energy to serve the shiftable loads throughout the day, and according to the market clearing price. The price regarding those two steps is defined according to the urgency of the agents to purchase energy, and is defined similarly to [11] and [24].

As to adapt the bids to the market prices, two constant variables, defining an action-pair, will modify the price of the two bid blocks regarding the flexible loads. As in [11] and similarly to [26], the two variables modify the price agents are willing to pay for energy according to their energy requirements.

By modifying these variables agents learn how to place bids optimally. To lead the agent on learning the optimal actions, a reinforcement learning algorithm will be introduced. The reinforcement learning algorithm applied in this work consists in the Q-learning algorithm, first developed by Watkins in [27].

The bids are aggregated by horizontal summation of all agents' bids at the intermediate level by the aggregator agent, as in [12]. After the bids aggregation, the market clearing will be performed as in [28] and [11]. The price is communicated to the aggregator agents, who will then transmit the information to the individual agents. The quantity of energy purchased by each agent is defined according to the market clearing price and the previously defined demand bids.

By the usage of the Q-learning algorithm, agents try a different set of actions a sufficient number of times, choosing the ones that minimize the cost of purchasing energy optimally. The usage of a ε-greedy policy, such as in [29], serves to guide agents as to either exploit or explore, leading them to firstable try agents randomly, and afterwards to exploit the most profitable actions. In order to test the efficiency of the model developed in this thesis, the results obtained throughout the usage of the developed model will be compared to the ones obtained in an approach without demand response, and the results obtained with a centralized scheme. Three different levels of DR penetration will be considered, as to observe the evolution of the impact of the shiftable side of the demand in the prices, traded volumes and consumption profiles.

In which concerns to the expected results, as in the electric vehicles example presented in [11], it is expected for the costs of purchasing energy from the new decentralized model of this work to be similar to the ones obtained with a centralized model, slightly higher, and considerably better than the results obtained with an approach without DR.

Thus, the importance of this work resides on creating an alternative to the centralized models, solving problems such as scalability and privacy and to the decentralized models relying on unidirectional communication, that don't consider the impact of the DR bids in prices, leading to undesirable outcomes such as load synchronization problems. The new decentralized model developed in this work is not only modelized as to reduce the cost of buying energy, but also to lead to a more uniform load diagram, avoiding high levels of consumption during peak hours.

Chapter 2

State-of-the-Art

Since the technologies related with power systems are developing exponentially, such as the possibility of the end-users to use smart meters and participate in DR programs, the management of the electrical networks will face many challenges, introducing new elements to the networks, such as the electric vehicles, photovoltaic and wind power.

The traditional top-to-bottom operation will be replaced by bidirectional communication, making it possible to the end-users not only to send information to the network, but also to sell power to the grid, creating new scenarios that will require the development of new approaches to deal with the crescendo in traded information.

In the traditional approach, demand was considered to be inelastic, and the supply was dispatched to match the demand. End-users would consume energy regardless the market prices, by submitting a quantity without putting a maximum price on it, telling the market that they consider the value of electricity to be infinite, as mentioned in [26]. This type of behavior would lead to a huge amount of demanded quantity of energy during peak hours, associated with higher costs of electricity usage.

To overcome those problems, a new operation model will be required for dealing with all the new features introduced in smart grids, such as the connection with decentralized demand-supply users. This operation model will consider recent features, such as the possibility of end-users to have access to demand response programs, leading to a more efficient usage of the grid and generation assets.

As for the infrastructure to be controlled and managed in all its complexity, an intelligent system will be required to match demand and supply in real time, such as the MAS.

The MAS meet these requirements, being its coordination alongside a MBC scheme well suitable for allowing the different agents to optimize their decision-making process.

Individual bids are defined at the lower level, sent to the intermediate level as to be aggregated, being the market clearing performed at the upper level. Actions are defined and tested throughout the usage of the Q-learning algorithm.

This chapter serves to introduce the state-of-art related with this work, presenting the theoretical development to achieve the expected results and the associated methodologies.

2.1. MAS

Multi-Agent Systems allow the management of complex systems, modeled as groups of intelligent agents capable of interacting with their environment, being also capable to adapt their behaviors accordingly.

Thus, MAS can be used in a huge variety of fields, such as market simulation, network control, or automation. An example of the usage of MAS in electrical distribution in net controls is presented in [30]. According to [30], a multi-agent system is a system that considers two or more agents. There is no overall system goal, only the local goals of the different agents are considered.

The three concepts that Wooldridge uses to define MAS are agent, environment and autonomy. Agents are defined as merely software, or even hardware, that can be placed in a specific environment, reacting to its changes autonomously. Their function is to represent an entity, such as the DR end-users.

Environment is defined as all the external elements to an agent, allowing the agent to modify it throughout their interaction. The environment shall be measurable as for the agents to be able to be aware of the impact of their action in the environment.

The concept of autonomy is related to the capability of the agents to adapt and answer to different environment states. Wooldridge also states a differentiation between agents and intelligent agents. Meanwhile both different types of agents have the capability to react to changes in its environment, intelligent agents also have not only the capability of being pro-active, presenting goal directed communication, but also social ability, i. e., capability to interact with other intelligent agents. In [21] there is another example of the usage of MAS in electrical networks. It is referred that even though MAS are highly complex, with a big level of intelligence, the complexity of the agents won’t evolve to a complex level as well. It is also referred that a way to make the process more efficient is to interact with an electricity market. Thus, a Market-Based Control scheme will be presented.

2.2. Market-Based Control

MBC is considered as a control strategy based on the micro-economy demand-supply model for price determination.

Regarding this theory, the different agents of the system compete for resources by allocating bids to the market, being the assignment of the resource determined by the market clearing price.

In [21], it is presented a scenario in which multiple agents behave differently from other agents, being the agents of this model defined as devices, such as shiftable, stochastic and storage devices. In this papers' approach, it is shown the combination between MAS and MBC as to control the multiple agents In which concerns to the scalability problem related with the centralized approaches, MBC is demonstrated to lead to a solution regarding that aspect, as MBC has low communication requirements. While the market clearing price remains as the control signal, information such as the agents' bids is defined and grouped as to optimize the cost reduction.

Reference [31] presents a market-based mechanism for scheduling short term energy consumption. It is presented a scheme considering an aggregator agent to be the intermediate between the individual agents and the control market. In this approach, a high number of distributed agents serves to justify the usage of a market-based control. The management of the decentralized resources is proven to be allowed by the MBC presented in this paper.

2.2.3. Privacy

As mentioned in [11], one important issue to consider about bi-directional control is the privacy issue. In bi-directional control schemes such as the centralized control, end-users need to share private information such as their load profiles, leading the central agent to have access to their personal information.

In the bi-directional control scheme developed in this work, the only information sent from the consumers is their individual demand bids, representing the quantity of energy they want to purchase and the price they are willing to pay for that amount of energy. The bids don't contain personal information, and privacy is no longer an issue.

2.3. Market-Based Control Scheme

Concepts such as MAS and MBC were presented on the last subsections, being referred that the interaction of both would lead to a well-established coordination of the agents of the system, providing a good control strategy to manage a highly complex system. This chapter serves to introduce a description of the appliance of both concepts in a scenario in which agents have the possibility to participate in DR programs.

In this subsection, the control scheme developed in the present work is presented, as well as the different agents, considering the concepts of MAS and MBC. The MAS was modeled within three layouts or levels. The upper level is represented by the electricity market. The intermediate level is represented by the aggregator agent, whose function is to act as an intermediate between the lower level and the upper level. The lower level consists on all the end-users that can participate in the DR programs. A scheme representing the three different levels is presented on Figure 2.1.

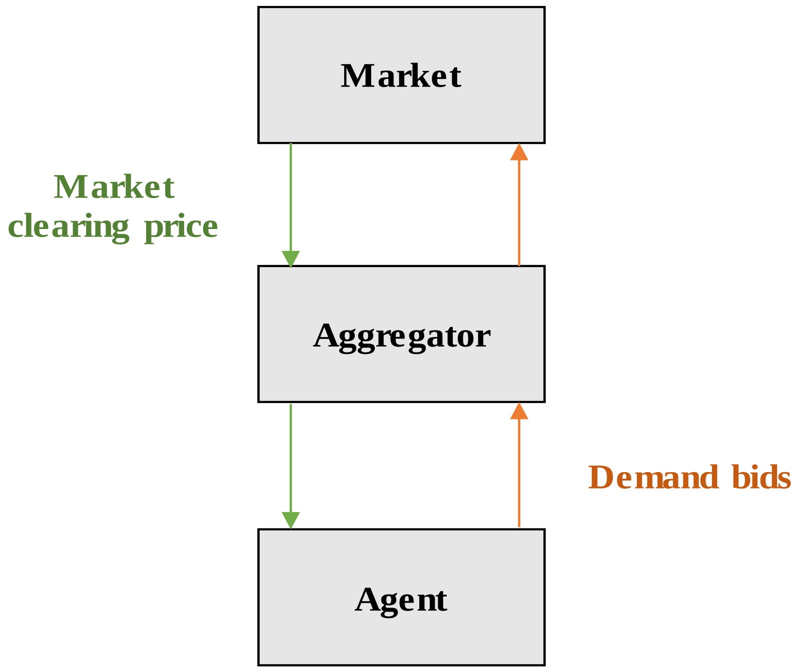

Figure 2.1 - Communication between the MAS levels

MBC is defined as a control strategy based on the micro-economy theory of demand-supply. According to this theory and the concept of MAS, the agents of the system compete for resources by setting their bids according to their willingness to buy energy. Those bids are aggregated at the intermediate level. Afterwards, the aggregated bids are communicated to the higher level, where the market clearing is performed.

The market clearing will lead to the determination of the market clearing price, that will be communicated in the opposite direction, i.e., to the lower level. The market equilibrium price will serve as the control signal for the end-users, whose amount of energy to buy will be set accordingly to the equilibrium price and their individual demand bids.

2.4. Load profiles

Since the scheme of this work is directed to residential loads, those household loads will be classified into two categories: controllable loads and critical loads, similarly to [25].

Critical loads are the loads that cannot be controlled, i.e., loads that must mandatorily be served at every time-step, otherwise major impacts will be observed in the consumer's lifestyle. Loads such as lightning and refrigeration are part of this type of loads.

On the other side, controllable loads consist in loads that can be controlled without major impacts on the end-user's lifestyle. Loads such as HVAC systems, clothes dryer, or water heatercan be served at several different times without causing much impact on the end-user's lifestyle. Thus, this type of loads can be seen as loads shiftable in time. Figure 2.2 presents a typical load profile to be used in this work.

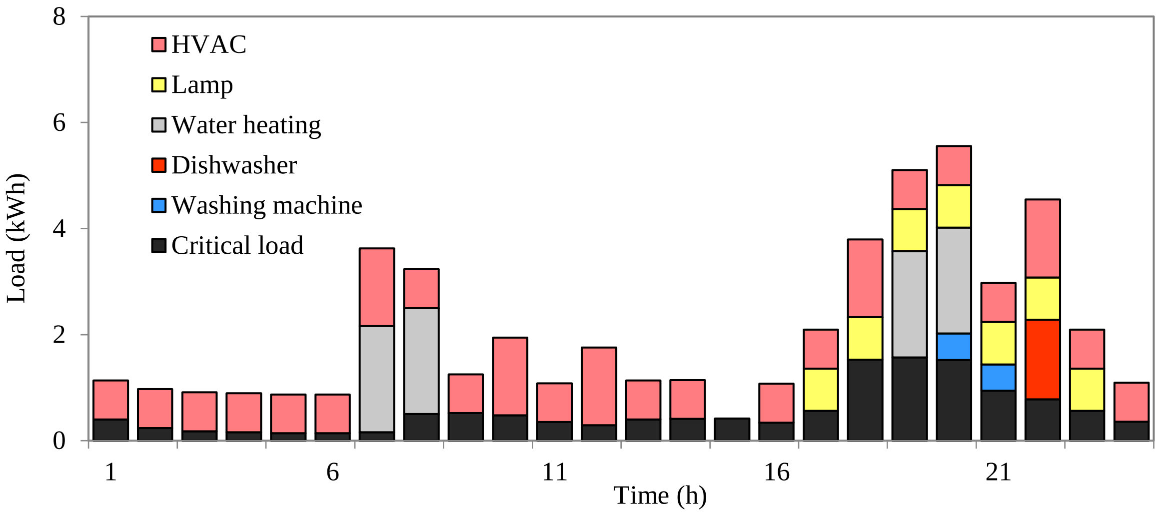

Figure 2.2 - Typical household consumption profile

In [25], a target control function for each house is defined, being adapted to this work and presented in below:

$$P_{shif} + P_C \leq P_L \tag{1}$$

$$P_{shif} = \sum_{i=1}^{N} P_{shif,i} \tag{2}$$

$$P_C = \sum_{i=1}^{N} P_{C,i} \tag{3}$$

Where:

- P_{shif} - sum of all controllable/shiftable loads *i;*
- P_C - sum of all critical loads i;
- P_L - maximum power a user can withdraw from the network.

Thus, the willingness of each agent to buy energy will have this formulation into account, considering that each agent needs to purchase at every given time step an amount of energy between the sum of all its critical loads at the current time-step and the maximum power they can withdraw from the net.

2.5. DR agents

The DR agents correspond to all the end-users participating on DR programs. Each agent has its own energy consumption pattern, in which both critical loads and shiftable loads are considered.

The agent's goal is to minimize their overall consumption cost by setting their bids accordingly to their urgency for buying energy. Throughout the cost optimization process, they need to learn how to place bids optimally as for them to compete with the other agents, taking advantage of the fact that a considerable part of their loads can be shifted in time.

2.5.1. Demand bids definition

As mentioned in [11], one of the major problems of the decentralized control schemes relying on unidirectional communication is that the impact of the DR bids in the prices are not taken into account, leading to problems such as load synchronization. Therefore, also similarly to [11] and [21], the load synchronization problems related to the decentralized control scheme relying on unidirectional communication are surpassed with the introduction of a bidding process.

This bidding process consists on defining bids for each individual agent, representing their willingness to buy energy, according to the market equilibrium price.

In [32] it is presented a descending temperature-price bidding curve considering demand response, in which the price to pay for electricity is related to the actual temperature of the household. In [26] a piecewise linear curve with ten different fixed price ranges is presented.

A three step piecewise linear curve is presented in [11] and [24]. The presented bid curve's first step is associated to the maximum price allowed in the market, being the other two steps associated to the urgency-dependent prices.

Even though the curve presented in [26] could lead to cost reduction, the piecewise linear curves presented in [11] and [24] not only are more urge-for-energy related but also raise the convergence speed of the learning code, as more blocks lead to a slower iterative process. Similarly to the bid curves in the mentioned papers, the bid function will be set in two different intervals.

The first one represents the inelastic demand interval, which in this work corresponds to the interval related to the critical loads.

The second interval corresponds to the flexible interval of the curve. This interval represents the willingness of the end-users to buy energy. It corresponds to a linear relation similar to the corner priority of the electric vehicles presented in [24]. In this interval, two different steps are considered, allowing two degrees of freedom.

The piecewise linear curve represented by three different steps is defined as in (4).

$$demand(price) = \begin{cases} p_{max}, & 0 \leq P \leq P_C \\ p_{state}, & P_C \leq P \leq P_{int} \\ p_{int}, & P_{int} \leq P \leq P_L \end{cases} \tag{4}$$

Both energy and price values defining the demand curve are based on state variables of the DR end-users' household, being those values related to the different hourly time steps. p_{max} corresponds to the maximum price allowed in the market, i.e., the maximum possible price to pay for energy.

The p_{state} variable, similarly to its definition in [11] and [24], represents the urgency of the DR end-users to buy energy. In these papers, p_{state} is defined as a fraction between the energy that needs to be purchased and the maximum possible amount of energy to buy before the departure of the vehicles. However, household DR end-users load profiles are considerably different, as for vehicles are able to store energy throughout the usage of batteries.

Meanwhile vehicles are able not to participate in the market in scenarios in which their battery is fully loaded, as an example, DR end-users have the necessity to participate in the market at every time-step, as for they do not only have critical loads to be served at every hour, being impossible to postpone the serving of those loads without causing major impacts on the end-users' lifestyle.

Vehicles have many different times ranges during which a certain amount of energy can be delivered, being the delivery of that energy flexible. The comprise of all the planning horizons may not be twenty-four hours, since vehicles may not need to participate in the market at every time step.

On the other side, household DR end-users' time ranges comprise the whole day, as for they have to participate in the market at every hour. In [26], a demand-type considering shiftable demand and its cost minimization formulation are presented.

Regarding this demand-type, the shiftable demand, which in this work is represented by the controllable loads, requires a certain quantity of electricity to be delivered within a given time range, being flexible regarding to the time of delivery within that range.

Consumers divide the planning horizon into multiple time ranges, with a restriction for the amount of energy to be delivered within that time range, i.e., the sum of the delivered electricity during the hours comprising the time range has to meet the total amount of energy to be bought within that time range.

In which concerts to this work, multiple time ranges could be considered. However, to simplify the bidding strategy, a single time range of twenty-four hours will be considered, making it necessary for all shiftable loads to be served within that time range. The formulation regarding the p_{state} is presented next.

$$p_{state} = \frac{t}{T} * \frac{P_C^t + P_{TS}^t}{P_L} \quad (5)$$

The time-step is defined by the variable t. The variable T represents the twenty-four hour time range. P_{TS} represents the total amount of shiftable loads to be served throughout the time range. The value of P_{TS} decreases as the agents buy energy. Thus, its value is also related to the time-step, being its value either equal or smaller when compared to its previous time-step value.

P_C defines the value the agent has to be connected at to satisfy all its critical loads at the current time-step. P_L represents the maximum power an agent can withdraw from the network. The goal of the formulation of p_{state} is to create a price to measure the urgency of the agent to buy energy.

The time fraction included in the formulation induces that as the end of the day is approaching, buying energy becomes more urgent, as there is less time to fulfil all shiftable loads. The variable P_C is included in the formulation to distinguish agents that at a certain time-step want to buy the same amount of shiftable energy, but have different willingness in which concerns to buy energy to satisfy the critical loads.

Thus, the higher the amount of energy an agent wants to buy, the higher they set their p_{state}. Also, as the end of the day approaches, the bigger their willingness to buy energy is.

Regarding the other variables of the demand formulation, p_{int} is an intermediate point between p_{state} and 0. P_{int} is defined as in below:

$$P_{int}^t = \frac{(P_L^t + P_C^t)}{2} \quad (6)$$

This approach for defining the demand bid curve and its variables is similar to the one in [1]. Thus, the only values for the agents to determine to their bids are P_{TS} and p_{state}, being all of the other values dependable on these ones. The demand bidding curve considering all these variables is presented in Figure 2.3. Considering the different types of demand bidding curves, two distinguished curves were cons idered regarding the status of the household end-users. The first type consists on consumers that have more urgency on buying energy, i.e., end-users that still have shiftable loads to serve. The typical bid curve for this type of users in presented in Figure 2.3.

The second type consists on users that already bought enough energy throughout the day and have no need to buy energy to satisfy the controllable loads. Thus, a single block bid curve will be defined for these type of consumers, being the value of p_{state} and consequently p_{int} equal to zero. In Figure 2.4, a typical bid curve for this type of costumers is presented. As a resume, the demand bid curves of the different users depend on their status variables. These status variables will define how their bid will impact in the market price, and according to the clearing price of the market, how much energy each agent will purchase.

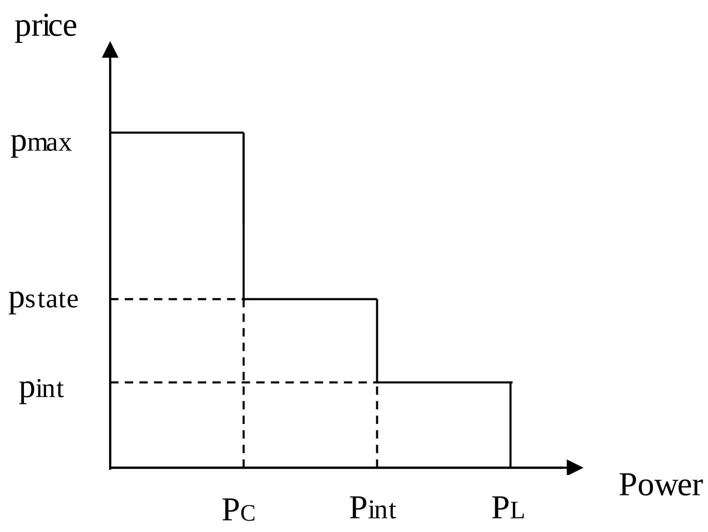

Figure 2.3 - Demand bid curve

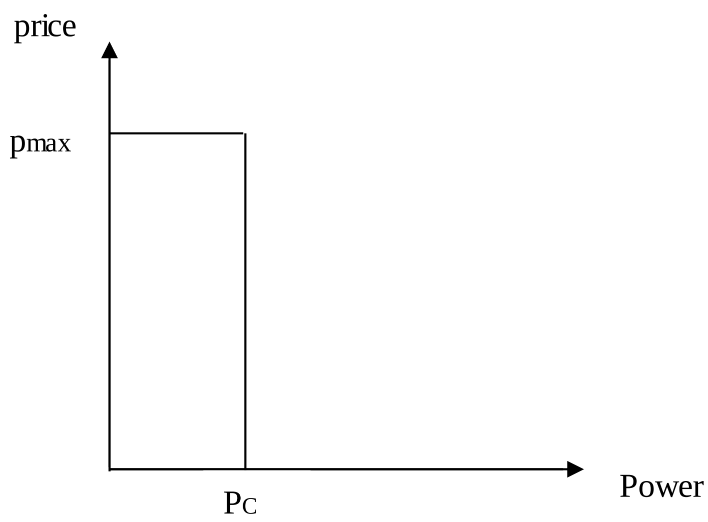

Figure 2.4 - Demand bid curve, no shiftable load

2.5.2. Agents' actions

For the DR end-users to adapt their bids to the price in the market, a set of actions must be taken in order to measure the sensitivity of their energy requirements and adapt it into a price. In [33], a pricing function in which the cost of purchasing energy is dependent on the level of consumption is presented. The price function takes into account the effect of the customers' consumption level so the customers will have a multiple price rating according to their consumption level. The formulation considers two different positive constant variables. One of the variables is independent of any other variables, and the second one is multiplied to the customers' consumption and their consumption level.

A similar approach that is even more adaptable to this work is presented in [11]. Similarly to the previous approach, the formulation in this paper also considers two constants, being the second one multiplied to p_{state}. Thus, the value of p_{state} is adapted considering an independent constant variable, and another constant multiplied to p_{state}. The adapted formulation is presented in below.

$$p_{state} = B_1 + B_2 * \left(\frac{t}{T} * \frac{P_C^t + P_{TS}^t}{P_L} \right) \tag{7}$$

In this scenario, B_1 represents the agents' interest on buying energy, i.e., the price that agents consider to be low enough to purchase energy even if there is no urge to buy it. B_2 represents the sensitivity to the urge of buying energy of the willingness to buy electricity. Meanwhile the state variables are energy-requirement dependent, actions may have a wider range of values previously defined. Changes in the values of the actions will lead to the optimization of the energy costs. By modifying the values of B_1 and B_2, agents will learn how to place bids optimally as to reach their cost minimization goal. For that, a machine learning algorithm is developed as to automatically test a wide number of actions and determine which actions lead to optimal solutions for minimizing the cost on purchasing electricity.

2.6. Aggregator agent

The aggregator agent represented at the intermediate level is a key element to integrate the DR bids within the rest of the bids, as mentioned in [34]. The quality of the aggregation process will determine the quality of the optimization process. In [35], a framework including the role of an aggregator agent is presented. In this framework, the function of the aggregator agent is to group electrical vehicles while trying to satisfy the energy requirements of the fleet. In [21], it is presented a scenario in which the role of the aggregator agent is likewise the role of a concentrator agent. The aggregator receives and concentrate the bids, sending them to the market afterwards. The clearing price of the market is sent to the end-users by the aggregator agent, leading the agents to buy energy depending on their individual bids and the clearing price of the market. This is the approach used in the work developed in this thesis.

2.6.1. Aggregation process

In order to optimize the cost of buying energy for the agents of the fleet, the aggregator has to aggregate the demand bids in a proper way as to not only meet the requirements of the fleet, but also to avoid misleading information to be sent to the upper level. The aggregated demand curve must be dimensioned in a way that approaches the costs of buying energy to the cost obtained with the centralized model to be presented in section 2.9.

As in [36], the aggregated function is a piecewise linear function. The aggregated demand is determined by the horizontal summation of the many different individual demand curves and has a number of intervals equal to the number of different prices of the individual demand curves.

Figure 2.5 demonstrates the horizontal summation process of the individual demand bids. Considering a scenario with two demand curves as defined in Figure 2.3 in which $p_{state1} > p_{state2}$ and $p_{int1} > p_{int2}$, are horizontally summed, the resulting curve would be similar as in the presented figure. The horizontal summation is done as in [37].

2.7. Reinforcement learning: Q-learning formulation

This subsection serves as an introduction to the presentation of the different known machine learning algorithms, as well as the definition of the algorithm to use in order to lead agents to an optimal bidding policy.

In [38], three different machine learning algorithms are defined: supervised learning, unsupervised learning and reinforcement learning.

In supervised learning, agents learn from labeled data, being able to classify on beforehand the information that it is receiving. A training set will be at the agent disposal so the agent will be led to generate a mapping for inputs and outputs.

On the other side, unsupervised learning consists on the agent learning from unlabeled data. The agent receives a large body of information, being its task to detect information that deviates from the average information. The so-called reinforcement learning algorithm consists in a machine learning algorithm whose learning process consists on analyzing the reward signals obtained after performing a particular action, being those rewards used for building a policy that allows agents to optimally choose the best future option.

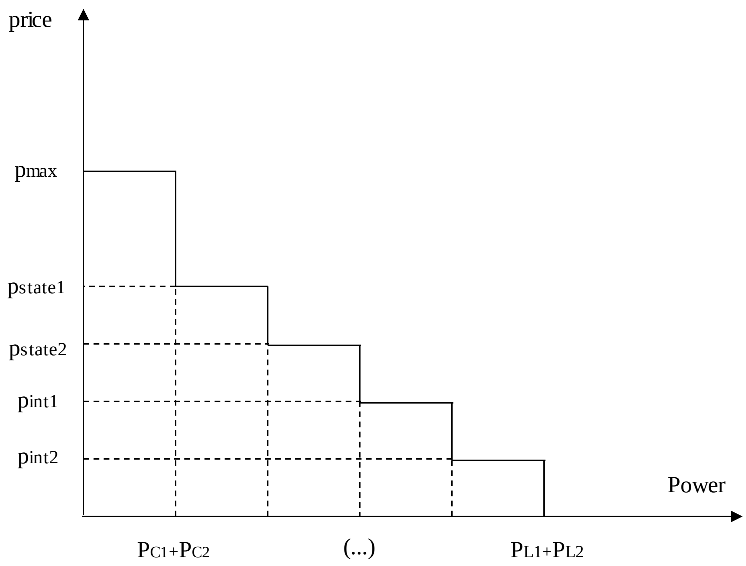

Figure 2.5 - Bid aggregation process

The goal of the reinforcement learning algorithm is to maximize the reward, i.e., to determine the sequence of actions that will lead to optimal rewards. Q-learning is an off-policy control algorithm, first defined by Watkins in [27]. By using this method, agents are able to perform different actions meanwhile they are in one of the different possible states. To every state-action pair, a reward will be assigned, being that assignment preserved in the Q-matrix and updated every time the same action is performed. The goal is to determine a strategy that will lead to the maximization of the different values of the Q-matrix, known as Q-values.

The usage of the Q-learning algorithm is justified due to the fact that, as mentioned in [39], the learned action-value function Q converges to its optimal value despite the followed policy. Thus, the analysis of the algorithm will be considerably simplified. Even though a policy is still being used and has effects on determining the state-action pairs to be visited, only the continuity of the update and visit of the different pairs is required for the method to converge.

The formulation and the algorithm of the Q-learning method will be presented in this chapter, as well a policy to lead agents on deciding how to choose the action to be performed, which in this work is the ε-Greedy policy. In Chapter 4, the implementation of the algorithm is presented, as well as the adaptation of the method to the work developed in this thesis.

2.7.1. Q-Learning algorithm

In [39], a description of the iterative process for updating Q-values is presented. Regarding this process, the designer firstable chooses the Q-values that correspond to each state-action pair. Every time an agent performs a particular action, its correspondent Q-value are updated.

The Q-learning algorithm is presented as following:

1) Initialize Q-matrix;

2) For each episode:

2.1) Choose a random initial state S;

2.2) For each episode:

2.2.1) Choose an action A from S using a policy derived from Q (for example, ε-greedy policy);

2.2.2) Take action A and observe the reward R and the new state S';

2.2.3) Update Q-matrix according to its formulation;

2.2.4) Set the state to the new state, repeating the process until reaching the final state.

The formulation for updating the Q-matrix will be presented in the next subchapter.

2.7.2 Q-learning formulation

Regarding the Q-matrix updating equation presented in [39], the update of the Q-values is done according to the following equation:

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha_t(s_t, a_t) * \left[R_{t+1} + \gamma * \max_a Q_a(s_{t+1}, a) - Q_t(s_t, a_t)\right] \quad (8)$$

where:

- $Q_{t+1}(s_t, a_t)$ - new Q-value, i.e., the updated Q-value;
- $Q_t(s_t, a_t)$ - old Q-value;
- $\alpha(s_t, a_t)$ - learning rate, varying from 0 to 1 and determining the weight of new information versus old information, and thus determining the convergence speed of the method;
- R_{t+1} - reward obtained for performing action *a*;
- γ- discount factor, that is a weighting coefficient of future or actual rewards;
- $\max_a Q_a(s_{t+1}, a)$ - estimated future values.
-

2.7.3. ε-greedy policy

Regarding the fact that the Q-learning algorithm provides the agent information for it to know which, are the most profitable actions, a policy for decision-making is required. This is justified by the fact that there is no decision rule associated with the process on beforehand.

By using a Greedy policy, agents will exploit the available information, always choosing the actions that maximize the reward. Nevertheless, by using a regular Greedy policy, agents will always try the most promising actions, i.e., actions whose reward values are the highest, instead of exploring the remaining options.

During the first iterations of a Q-learning algorithm, all the actions are unexplored. The first actions to be taken by the agents will lead their correspondent Q-values to be updated more times than the other remaining values. Thus, by using a greedy policy, the first actions to be taken will be over-explored, and the agent will not learn any new information from the unexplored actions.

Therefore, a question regarding the concept of Exploration or Exploitation surges. Up to a certain point, agents will have to estimate if there is a need to explore more actions, or if it is time to exploit the actions that led to the best results so far.

The agent has to decide between Exploration or Exploitation by measuring the certainty of being at the best possible state. If not, actions should keep being explored as for their respective Q-values to be optimized.

In [29], the usage of a ε-Greedy policy is presented. By using this policy, agents will choose with probability ε a random available action a∈A, and with a probability of 1-ε the action that maximizes the expected reward.

According to [29], the main reason to adopt a policy that sometimes selects an action believed not to lead to the highest expected reward is that by taking all the actions a sufficient number of times, agents are able to correctly identify which actions lead to their expected reward.

As soon as the agents suppose that all the actions were taken a sufficient number of times, agents may then start to exploit the highest-reward' associated actions. An algorithm considering the ε-Greedy policy is also presented in [29] regarding the usage of the mentioned policy in a multi-agent system. The implementation of the algorithm, as well as its adaptation to the work developed in this thesis, is presented in Chapter 3.

2.8. Market Clearing

Regarding the known decentralized schemes, agents need to receive from the upper level, i.e., the market, a control signal that will lead them to act in terms of buying energy.

In a Market-Based Control strategy, the control signal is the market clearing price. This price is determined by the microeconomic theory of demand and supply. Differently from decentralized control schemes relying on unidirectional communication, the DR end-users' bids will have an impact on the market price. After the market price is determined, agents will respond to it according to the bids they previously set.

The price will be determined with a simple market clearing, presented in the next subchapter.

2.8.1. Electricity Market

As for the market operator to determine the market price, it has to match supply and demand bids, i.e., to determine a price to pay for energy that satisfies both demand and supply in term of quantity of energy to be satisfied/produced. The market operator's goal is to maximize the social welfare, which according to the microeconomy theory happens when the consumer and producer surplus are maximized.

2.8.2. Market Clearing formulation

In [28], an optimization-based approach is presented and proposed for clearing the market in a smart grid. This approach's formulation considers four different terms to be maximized, being those terms the utility for suppliers, consumers, bilateral contracts and electric vehicles aggregators.

To every one of the terms an inequality equation is associated. An equality equation is also present in the formulation, representing the balance of supply and demand in the grid. The work to be developed in this thesis doesn't consider bilateral contracts. Also, it doesn't consider the introduction of electric vehicles in the grid as well. However, electric vehicles behavior in the control scheme is similar to DR end-users' behavior. Thus, a similar term to the one representing the electric vehicles will be considered in the market clearing formulation.

Regarding the updates to be done to the formulation in [28], it is observable in [11] a formulation for clearing the market that not only is similar to the pretended one, but also applied in a similar Market-Based scheme. The formulation is presented in below.

$$\max_{P_c^t, P_a^t, P_p^t} \left(\sum_c p_c^t P_c^t + \sum_a p_a^t P_a^t - \sum_p p_p^t P_p^t \right) \tag{9}$$

subject to

$$0 \le P_l^t \le P_{max,l}^t \quad \forall l = \{c, a, p\} \tag{10}$$

$$\sum_c P_c^t + \sum_a P_a^t = \sum_p P_p^t \tag{11}$$

where:

- P_c - accepted demand;
- P_a - aggregated DR bids;
- P_p - supply;
- p_c, p_a and p_p – price.

The objective function to be maximized ensures that social welfare is maximized.

Equation (10) is the inequality restriction that will ensure that the values of energy to be produced and served are within the acceptable limits.

Equation (11) is the equality equation to ensure the balance between demand and supply.

In Figure 2.6 the resolution of the market clearing problem is presented visually. The descending curve represents the demand, while the ascending curve represents the supply. Maximizing the social welfare consists on maximizing the area between the demand and supply curves, i.e., determining the intersection between orange and blue curves. The intersection point is defined by the quantity to be produced/purchased in every time step and the market clearing price.

The market clearing price is determined as in [40].

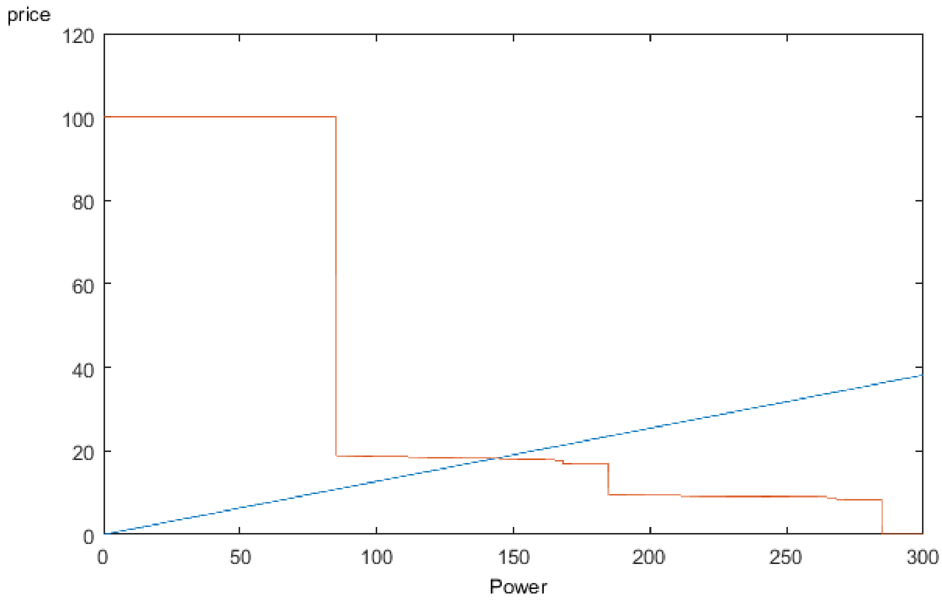

Figure 2.6 - Market clearing example

2.9. Centralized approach

In order to determine the effectiveness of the new decentralized scheme proposed in this work, the results are compared to the ones obtained in a centralized market. In the framework pursued in this work, individual agents directly participate in the market by submitting their bids and respond to the market clearing price according to them. On the other side, in a centralized model, bidding is managed centrally by an aggregator.

In a centralized model, aggregators bid into the market on behalf of the fleet's customers. At the same time the aggregator meets the energy requirements of the fleet, it directly controls purchasing energy. Similarly to the decentralized approach, the goal is to minimize the costs of buying energy while satisfying the DR end-users' constraints. Those constraints are aggregated similarly to the constraints of a virtual battery as in [11]. The problem is modeled as a bi-level program, i.e., the aggregator's cost minimization problem represents the upper level, being the market clearing associated with the lower level, similarly to [11].

2.9.1. Formulation

The formulation representing the cost minimization problem associated with the upper level considers not only the minimization objective function, but also the constraints representing the aggregated constraints of all individual end-users, and is presented as following.

$$\min_{E_0,P_a^t} \sum_{t=1}^{24} p^t P_a^t \tag{12}$$

Subject to:

$$E_a^t = E_a^{t-1} + P_a^t \times \Delta t \tag{13}$$

$$E_a^0 = E_a^T \tag{14}$$

$$E_{min,a}^t \leq E_a^t \leq E_{max,a}^t \tag{15}$$

$$P_{min,a}^t \leq P_a^t \leq P_{max,a}^t \tag{16}$$

The objective function consists on minimizing the costs of purchasing energy throughout the day.

Equation (13) determines the evolution of the total quantity of energy to be satisfied during the whole day. Equation (14) guarantees that enough energy will be purchased during the day. Equations (15) and (16) respectively represent the energy and power boundaries. The parameters regarding the power and energy constraints are determined from the individual consumption patterns from the considered agents.

The lower level of the problem is represented by the market clearing problem presented in subsection 2.8.2, regarding equations (9) - (11).

The connection between the upper and the lower level is represented by the Lagrange multiplier of the equality equation (11).

Chapter 3

Implementation

This chapter serves to introduce the different steps regarding the implementation of the different simulations. The particular application of the formulation presented in previous chapters is adapted in this chapter to the work developed in this thesis. All simulations were implemented using *MATLAB* and the tool *Solver* from *Excel*.

3.1. Main implementation

The steps regarding the implementation of the simulations are presented in below:

1. Definition of the actions: As the day begins, the actions to be performed by each individual agent are defined according to the E-Greedy policy;
2. Definition of the bids: At the lower level, the demand bids representing the willingness of each individual agent to buy energy are defined;
3. Aggregation of the bids: At the intermediate level, the bids of all individual agents are aggregated;
4. Market clearing: The individual agents' bids are integrated with the inflexible demand and the supply, followed by the performance of the market clearing;
5. Updating state of consumption: Individual agents respond to the price signal according to their bids, and update their state of consumption accordingly. At the same time, the reward associated with the agent and the performed action is simultaneously updated;
6. Updating the Q-matrix: At the end of the day, the Q-matrix is updated. A new set of actions is chosen for each individual agent.

A diagram resuming all the steps is presented in Figure 3.1.

3.2. Q-learning implementation

The Q-learning formulation presented on section 2.7 considers agents to be in a specific state at every iteration. Actions are chosen according to the state in the previous iteration, and future rewards are taken into account when choosing the action with the highest associated reward.

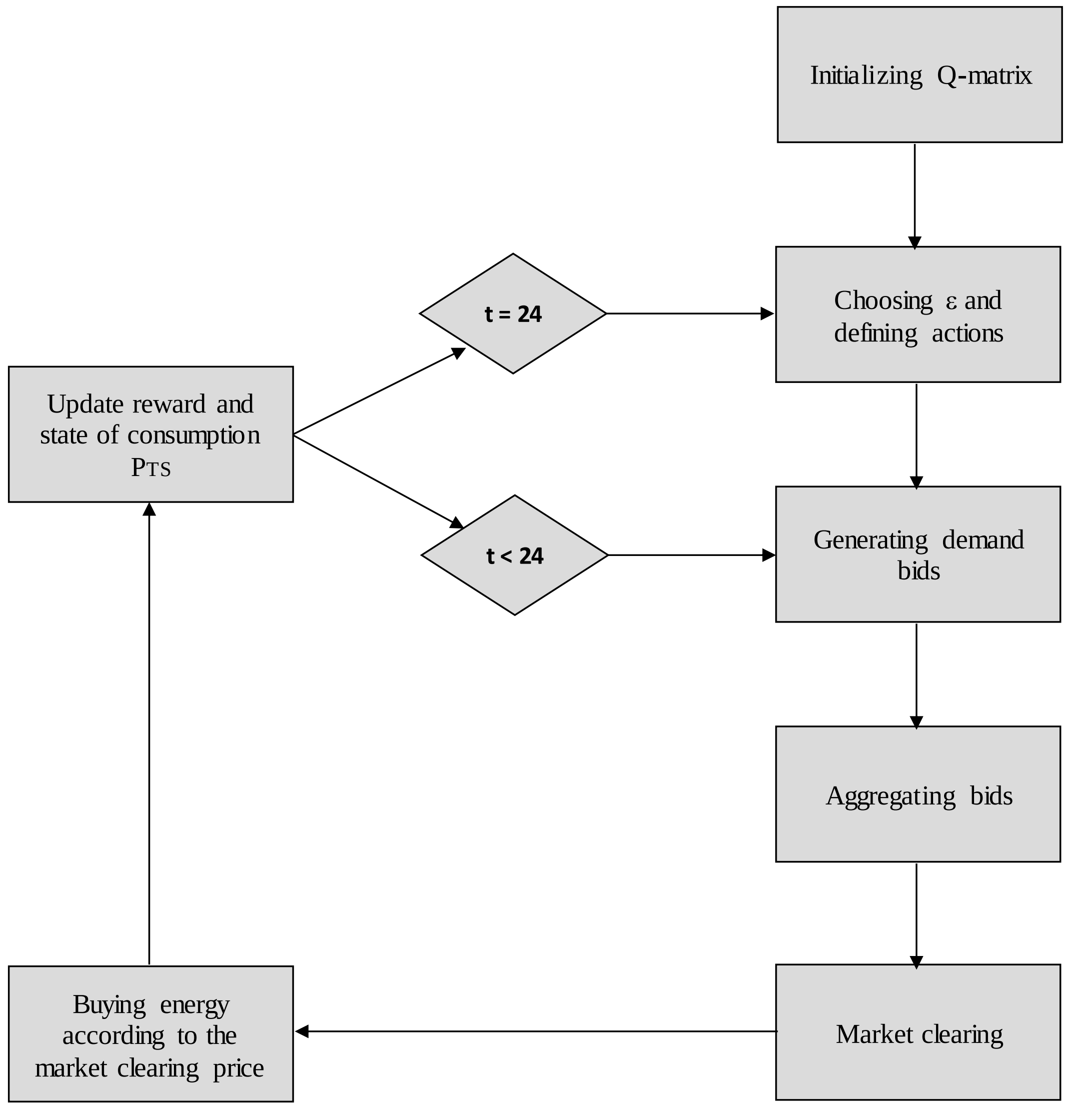

Figure 3.1 - Model implementation algorithm

However, the state of the agents is already considered in the bids definition, since the prices on the flexible demand bid blocks are defined according to the state of consumption. Thus, states are not considered in the Q-learning formulation.

The consequentialism of the non-consideration of the states in the Q-learning algorithm regards two consequences. The first one is that future rewards aren't considered. By not considering a sequence of states, rewards regarding that sequence will not exist. Thus, the algorithm's formulation is updated as following:

$$Q_{t+1}(a_t) = Q_t(a_t) + \alpha_t(a_t) * [R_{t+1} - Q_t(a_t)] \quad (17)$$

The second consequence is that rewards will be associated with the action performed by the agents in a particular iteration, and thus they are iteration-independent.

Therefore, the traditional Q-matrix can be updated as following.

$$Q = \begin{bmatrix} Q_{1,1} & \cdots & Q_{1,m} \\ \vdots & \ddots & \vdots \\ Q_{n,1} & \cdots & Q_{n,m} \end{bmatrix} \tag{18}$$

Meanwhile in a common Q-matrix the lines represent the states and the columns the set of possible actions, the new matrix can be resumed, for each agent, as a single line in which the columns keep representing the actions. In (18), *n* represents the set of agents, and *m* the set of possible actions. A similar simplification of the Q-learning algorithm in a MAS was already developed in [29] and [41].

3.2.1. Reward

The formulation for calculating the reward in presented in below:

$$R_n = -\left[\sum_{t=1}^{24} p^t\, P_n^t + w(|E_{exp} - E_{24}|)\right] \tag{19}$$

The first term represents the cost of purchasing energy throughout the day, being p^t the market clearing price in time-step t, and P_n^t the power agent n is connected at time-step t. The second term is a penalty for the agents who either fail to buy enough energy throughout the day or buy more than the expected amount of energy.

This term also leads the learning algorithm to a faster determination of the actions that maximize the rewards, as for actions that lead agents to buy less energy than expected can deceptively lead the algorithm to believe those actions are in fact the one with a higher associated reward. In (19), E_{exp} represents the expected amount of energy to be bought throughout the day, while E_{24} represents the energy that was in fact purchased by the agents during the day.

Another feature of the penalty term is to compensate the market for minor deviations in the bought quantity by each agent. Due to the nature of the demand bids, agents only have three steps of freedom when it is about purchasing energy. When they answer to the market according to their bids, agents are likely to purchase more or less energy than the accorded amount during the market clearing. In case the agents end up purchasing more than the accorded amount, a higher quantity will be set to a market clearing price that don’t correspond to the price. Thus, the penalty term compensates the market by the end of the day, as agents with deviations in the expected amount of purchased energy will pay according to the level of deviation.

3.2.2. Q-learning parameters

To optimize not only the cost minimization but also the convergence of the algorithm, it is important to model the learning rate α and the parameter ε associated with the ε-Greedy policy. During the initial iterations, in which most actions remain unexplored, a sufficiently high value is set to α. As the actions with higher associated rewards are determined, the value of α will decrease, because the urge to gather more information decreases as well.

In which concerns to the parameter ε, its value will determine the probability of exploring vs exploiting. Since it is essential for the algorithm to try all different actions a sufficient number of times, it is necessary to firstable explore over exploit. Thus, at the beginning of the iterative process, ε is set to a value high enough for the mentioned purpose. As the optimal solutions arise, the value of ε is reduced, as for exploitation to win over exploration, leading to the final optimal solution. A more extended analysis on determining the most suitable values for B_1 will be presented in Appendix A.

3.3. Critical and controllable loads definition

In 2.4. loads are divided into two categories, critical and controllable loads. Each load type is divided according to the nature of the loads constituting it, i.e., critical loads consist on loads such as refrigeration, not being possible for the end-users to post-pone their consumption, and controllable loads are defined as loads such as washing machines, being possible for the end-users to purchase energy to satisfy this type of loads whenever they wish.

Multiple levels for the demand bid curve presented in 2.5.1. could be considered, so the multiple agents could satisfy individual loads according to the market clearing price. However, as the goal of this work is to prove the efficiency of the model in reducing the costs of purchasing energy while solving problems associated with the centralized models, merely the summation of the loads is considered, i.e., demand is divided in two: the critical part of the demand, and the shiftable part of the demand. Agents buy enough energy to satisfy the critical load, and according to the market price they purchase energy to satisfy the shiftable loads.

3.4. Definition of the actions

A fundamental aspect of this work is to shape the demand bids in order to develop optimal bids so the many individual agents can diminish their energy purchasing costs.

Considering that the agents' bids are defined by a pair of actions to be chosen at the beginning of each day, it is important to define an adequate amount of actions so the agents may have not only a wide range of actions to opt for, but mostly to avoid non-optimal results.

According to the equation (20), the willingness to pay for a certain amount of energy is not only defined by their individual requirements, but also by the actions defined by the parameters' pair B_1/B_2.

$$p_{state} = B_1 + B_2 * \left(\frac{t}{T} * \frac{P_C^t + P_{TS}^t}{P_L}\right) \tag{20}$$

Regarding Figure 2.6, these parameters determine the impact of the aggregated bids blocks in the market clearing price. If the parameters are dimensioned to low values, agents might fail to purchase energy. If the parameters have high values, agents are led to buy high values of energy as the day begins, increasing the competition for resource during the first hours of the day, and therefore leading to a huge variation between maximum and minimum value of purchased energy throughout the day. Thus, an adequate approach on correctly defining the mentioned parameters will be presented as for the individual agents to achieve their goal.

3.4.1 Defining B_1

The parameter B_1 represents the agents' willingness to buy energy even if their requirements are considered to be low. According to [11], the higher the agent sets its value for the similar parameter b_v, the more likely it is to quickly purchase enough energy to satisfy the agents' needs, needs that in this work are represented by the amount of shiftable demand to satisfy throughout the day. Thus, defining B_1 is similar to defining B_2 in terms of impact in the optimality of the results.

However, a rule for defining values for B_1 is defined. Since B_1 is a constant value that serves to modify p_{state} and is not affected by a fraction, the parameter is defined taking into account the market clearing prices. Regarding the fact that the agents' goal is to diminish the costs of buying energy, it won't be willing to pay a higher value than the maximum clearing price of the market. Also, it is expected that the market clearing prices will raise during valley hours, as the goal of the individual agents is to purchase energy during those hours. Therefore, the minimum value to consider for B_1 should not be lower than the clearing price of the market, since that could lead the agents to fail on buying energy.

In Figure 3.2, it is possible to observe the clearing prices of the market in an approach without considering demand response. The clearing prices of the market in this scenario vary from 7.76 to 91.62 €/MWh. Thus, the range of values to attribute to B_1 must be between those values. A more extended analysis on determining the most suitable values for B_1 will be presented in Appendix A.

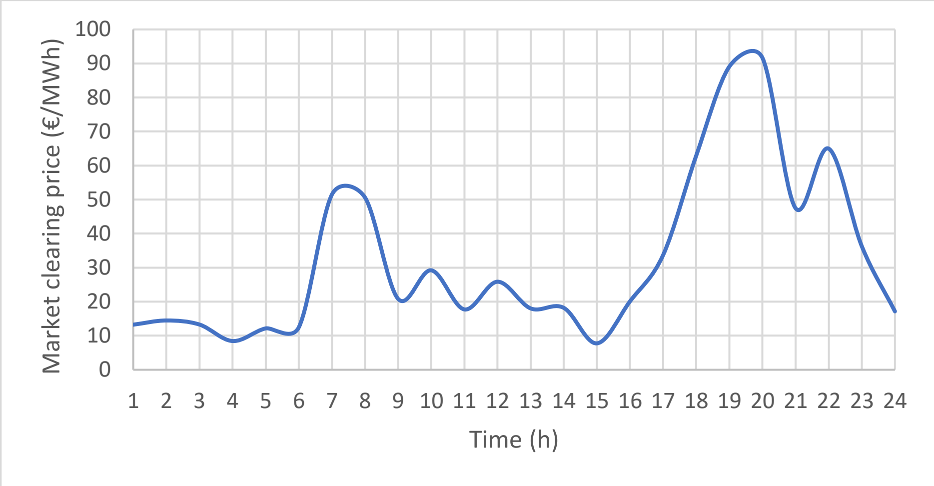

Figure 3.2 - Market clearing price without DR

3.4.2. Defining B2

According to its definition in Equation (20), B_2 serves to measure the sensitivity to the consumption urgency. Thus, properly attribute to this parameter an adequate range of values is highly important.

Low values of B_2 lead to an inadequate attributed importance to the agents' willingness to buy energy, leading them to fail on buying enough energy to satisfy their requirements.

High values of B_2 lead to high prices in the second and third bid blocks presented on Figure 2.3 in subsection 2.5.1., leading the agents to purchase an inadequate amount of energy considerably fast, pushing agents not only to non-optimal results, but also to high peaks of demand from DR end-users during certain time steps.

A more extended analysis on determining the most suitable values for B_2 will be presented in Appendix A.

Chapter 4

Results

The goal of this work is to model a decentralized market-based scheme considering demand response relying on bidirectional communication. To test the efficiency of the model, its results are compared to the ones obtained by modeling a centralized scheme and to the results obtained with an approach without demand response.

As to test the influence of the DR penetration level on the optimization of the results, three different DR penetration levels will be considered. The first one is 30 %, a situation in which DR starts having a noticeable impact not only in the market prices but also the load diagram of the fleet considering DR. The second one is 15 %, a situation expected to be observable in a near future. The third one is 60 %, a situation in which DR dominates over the inflexible demand part of the fleet.

A fleet of 100 loads is considered as to demonstrate the new decentralized model. Both demand and supply don't change from one iteration to another, as the goal is to test and define the actions that optimize the cost reduction.

4.1. 30% DR penetration

The first scenario tested in the base case considers a level of 30 % DR penetration. Regarding the DR level and the model developed in this work, the impact of the DR program in the prices, traded volumes and consumption profiles is observed and discussed in below, as well as the evolution of the average cost of purchasing energy, the evolution of the reward and also errors entailed with the model of this work.

- **Market Clearing Price:**

As a mean of interpretation of the changes in traded volumes and consumption profiles, the impact of introducing the DR program developed in this work to 30 % of the fleet in the market clearing prices is analyzed. Figure 4.1 presents the market clearing prices throughout the day both in the scenario without DR and the decentralized scheme developed in this thesis.

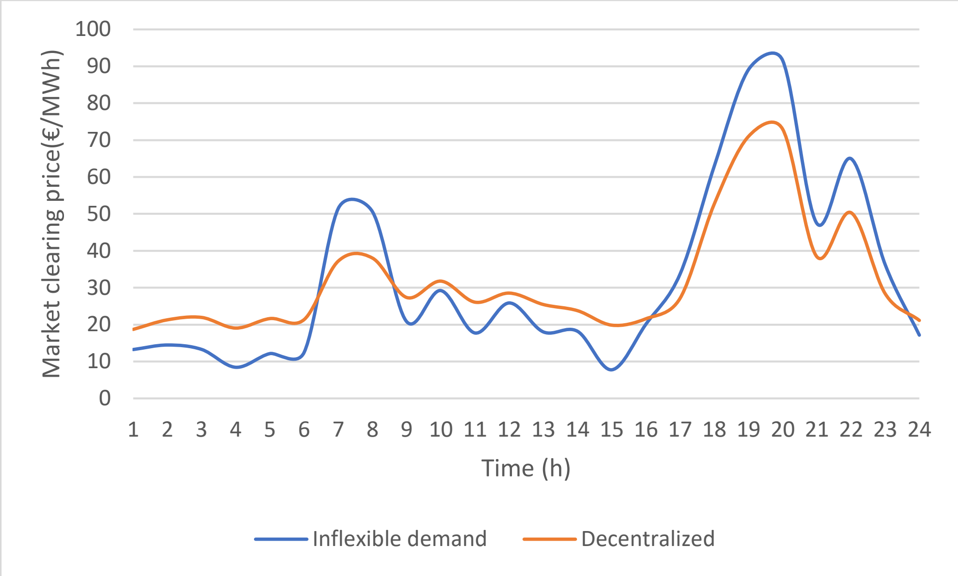

Figure 4.1 - Market clearing price, 30% DR penetration

As expected, the market clearing prices during valley hours raised, meanwhile during peak hours they diminished. The variation between the minimum clearing price of the market and the maximum value diminish when using a decentralized market-based scheme. This happens due to the fact that the market clearing price is influenced by the quantity of energy demanded at every time-step. DR programs lead end-users participating in the programs to demand more energy when the prices to pay for it is lower. Thus, DR end-users demand a higher quantity during valley hours, which will raise the clearing price of the market during those hours. Since end-users fulfil their energy requirements during valley hours, their demanded quantity during peak hours is lower, leading the quantity demanded, introduced with the rest of the inflexible demand of the fleet, to be lower, and therefore diminishing the price during peak hours.

Table 4.1 presents the variation between minimum and maximum clearing price in both approaches. As it can be observable, using the DR model developed in this work leads to a reduction of the variation between the maximum and the minimum of 35.28 %, when compared to an uncontrolled approach.

For further analysis, the next subsection presents the consumption profiles of the 30% of the fleet participating in the DR program developed in this work.

Table 4.1 - Variation between maximum and minimum MCP, 30 % DR penetration, in €/MWh

	Inflex. Demand	30 % DR
Min	7,76	18,70
Max	91,62	72,98
Max - Min	83,86	54,28
Variation (%)	35,28%	

- **Consumption profiles:**

Regarding the 30 end-users participating in the DR program developed in this work, the cost of purchasing energy without demand response and considering 30 % DR penetration are compared in Figure 4.2.

Figure 4.1 presents the change in the market clearing prices influenced by a variation in the demand throughout the day. As expected, it is possible to observe that DR end-users tend to purchase energy when it is cheaper.

DR agents buy energy according to their demand bids and the market clearing price. The first six hours of the day comprise the hours in which energy is cheaper. Thus, the willingness of the end-users to buy energy during these hours is higher than in the rest of the day, leading them to maximize the purchasing of energy during the first hours of the day, as it is possible to observe in Figure 4.2.

Table 4.2 presents the variation between minimum and maximum energy bought in both approaches. As it can be observable, using the DR model developed in this work leads to a reduction of the variation between the maximum and the minimum of 32.74 %, when compared to an uncontrolled approach.

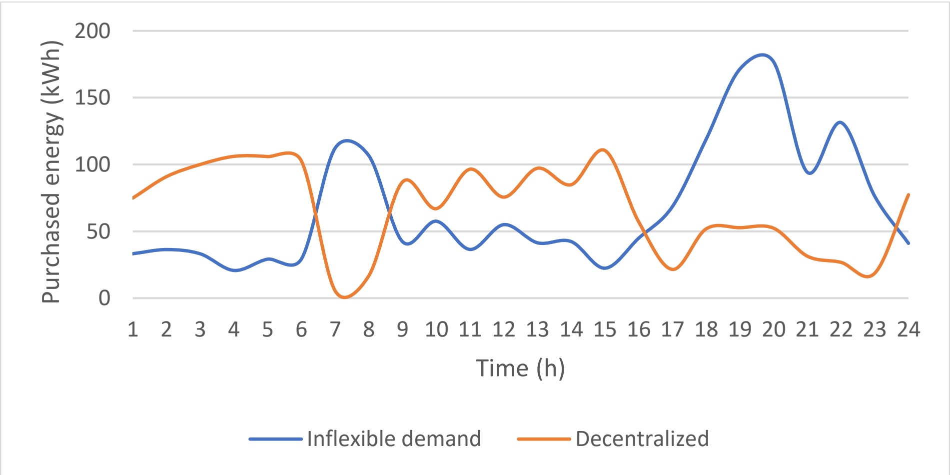

Figure 4.2 - Purchased energy, 30 % DR penetration

Table 4.2 - Variation between maximum and minimum quantity of bought energy, 30 % DR penetration, in kWh

	Inflex. Demand	30 % DR
Min	20,74	5,48
Max	176,86	110,49
Max - Min	156,12	105,01
Variation (%)	32,74%	

- **Traded volumes:**

Considering that one of the goals of introducing DR to end-users is to diminish the variation between the maximum and minimum traded volume, leading to less demand during peak hours, the impact of introducing 30 % of the fleet to the DR program of this work is analyzed. Figure 4.3 shows the traded volumes both without demand response and considering the scenario of this subsection. As it can be observable, the introduction of 30 % of the fleet to the DR program developed in this work reduces the overall consumption of energy during peak hours and raises the consumption during valley hours, as pretended.

Table 4.3 presents the variation between minimum and maximum energy bought in both approaches.

As it can be observable, without demand response, the traded volumes throughout the day vary from 56.34 kWh to 611.60 kWh, and with the scheme developed in this work, the traded volumes vary from 142.57 kWh to 487.11 kWh. Thus, the model developed in this work demonstrates to be effective in which concerns to diminish the variation between maximum and minimum value of traded volumes, as the variation using the DR model of this work leads to a reduction of 37.95 % between maximum and minimum when compared to an uncontrolled approach.

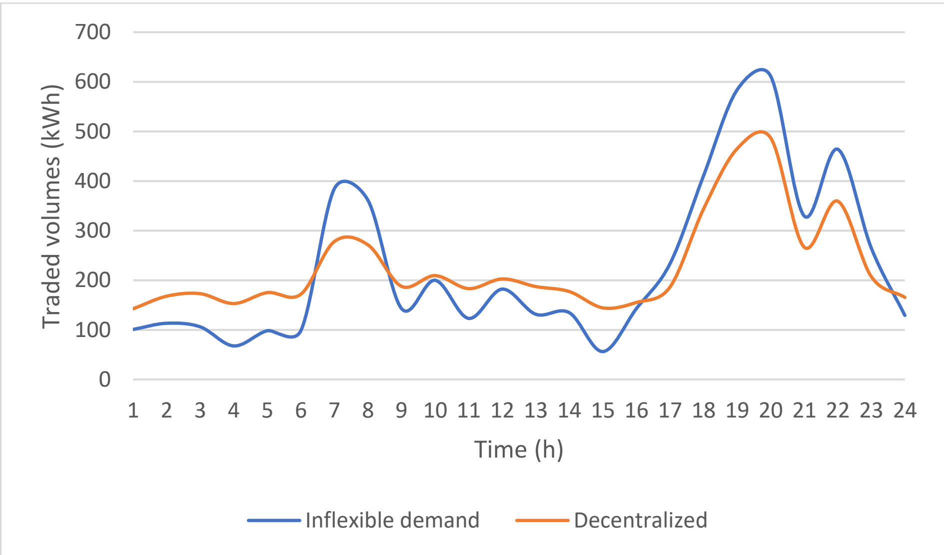

Figure 4.3 - Traded volumes, 30% DR penetration

Table 4.3 - Variation between maximum and minimum traded volumes, 30 % DR penetration, in kWh

	Inflex. Demand	30 % DR
Min	56,34	142,58
Max	611,60	487,11
Max - Min	555,26	344,53
Variation (%)	37,95%	

- **Costs of purchasing energy:**

The average cost of the purchased energy by the agents is calculated throughout the following formulation:

$$\frac{\sum_1^n \sum_1^{24} price^t * Power^t}{\sum_1^n \sum_1^{24} Power^t} \tag{21}$$

As a benchmark to compare the effectiveness of the model developed in this work, the average cost will be compared to the average cost in the approach without demand response. In Figure 8.4 it is possible to observe the average cost of purchasing energy throughout the iterative process.

The cost tends to vary considerably while the agents are still testing the many possible actions. As the agents start identifying the optimal actions, the costs of purchasing energy diminish considerably to an amount of approximately 29.75 €/MWh. In which concerns to the average cost of buying energy in a scenario without demand response, the average cost as a constant value of 48.37 €/MWh. Thus, it can be concluded that the DR model developed in this work leads to a considerable cost reduction, especially when agents determine the most profitable actions.

- **Error:**

Since the different agents of the fleet purchase energy taking into account their previously defined demand bids and the market clearing price, an error entailed with that method is expected, as it can be seen in Figure 4.5.

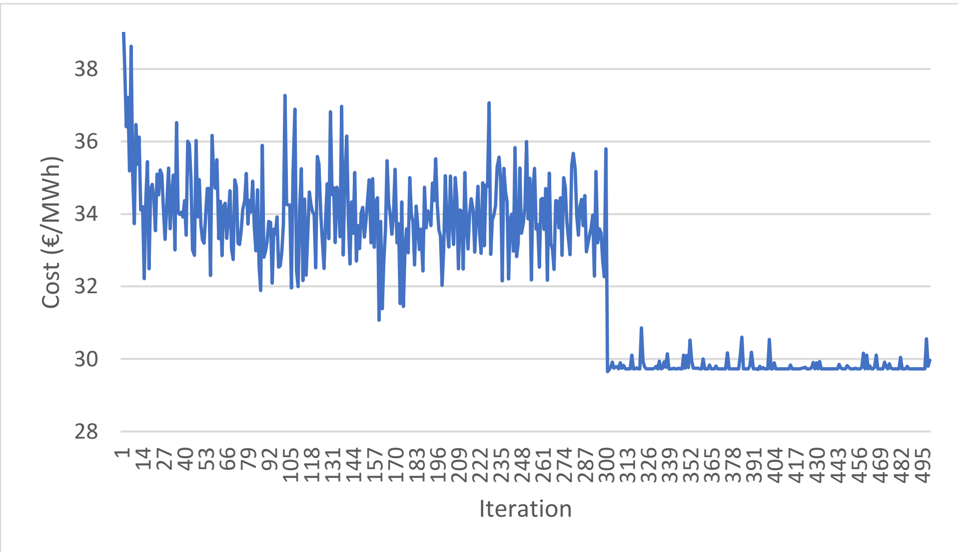

Figure 4.4 - Average cost of purchasing energy

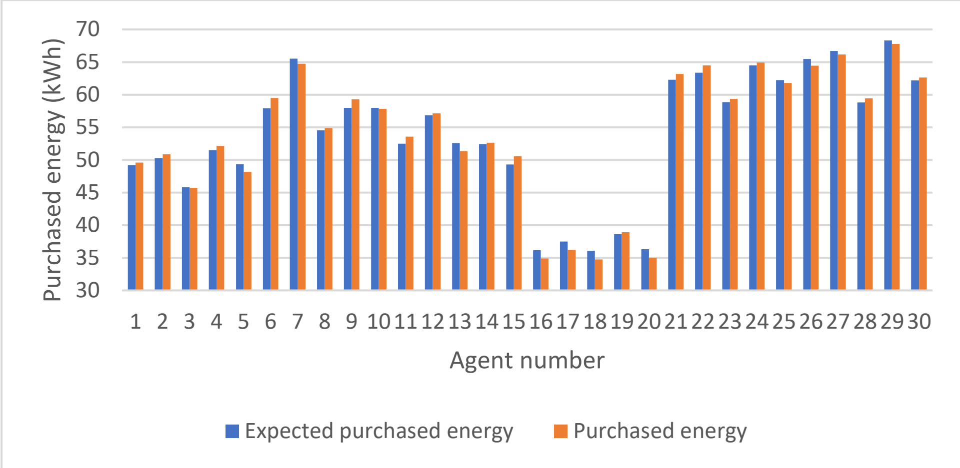

Figure 4.5 - Cost of purchasing energy

As it can be observable, the variance between the expected value of energy to purchase and the real value is not severely significant. Figure 4.6 show the variation between E_{exp} and E_{24} in terms of percentage. As it can be seen, during some hours of the day, the variation between both values is considerable, such as in hour 6. Plus, analyzing the variation it is observable that most of the times the agents of the fleet purchase more energy than the expected quantity, leading the market to sell energy cheaper than the accorded amount, as agents purchase a superior quantity than previously defined. However, the penalty term established in Chapter 3 in Equation (18) is used so the different agents of the fleet can compensate the market either when they fail to purchase energy or when they purchase more than the established amount.

Another feature of the penalty term is its optimal solution leading aspect, i.e., solutions in which agents buy less energy than it was supposed or buy more for an unfair price to the market will lead to smaller rewards, leading the agent to prefer actions with the minimum possible deviation from the expected amount of energy to purchase.

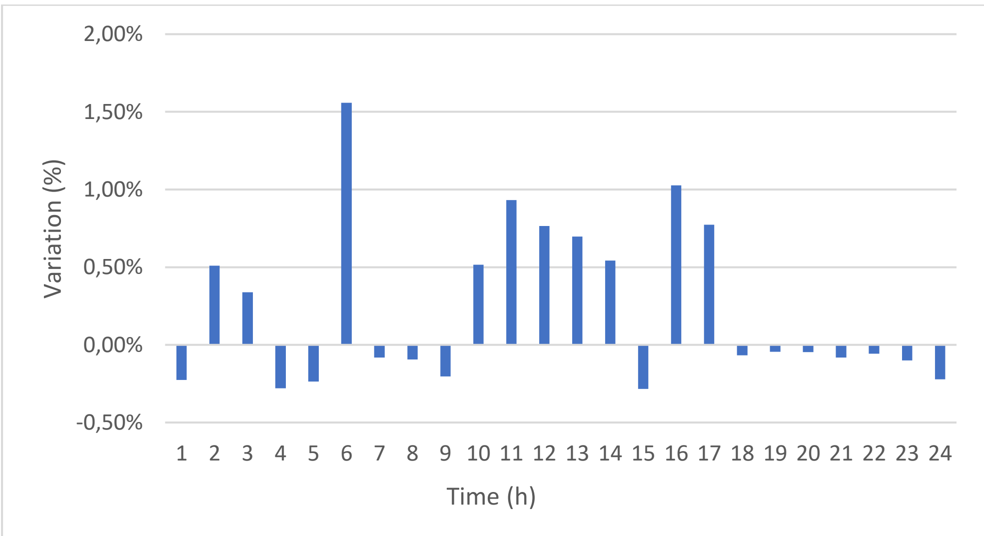

Figure 4.6 - Variation between assigned quantity of energy and purchased quantity of energy

- **Evolution of the reward:**

The evolution of the reward throughout the iterative process will determine the optimality of the obtained results, as a lack of convergence in the iterative process lead to sub-optimal results.

Figure 4.7 demonstrates the evolution of the fleet's reward throughout the iterative process. During the first iterations of the process, agents test new actions to learn which actions might lead them to higher rewards. As the agents starts to learn which actions are more profitable, the agents start choosing the actions that demonstrate to possibly be the optimal actions. By the end of the iterative process, agents seem to have reached the optimal actions, as the value of the total reward of the fleet tends to converge.

4.1.1. Centralized vs. Decentralized scheme

Considering that one of the goals of this work is to validate the developed decentralized scheme as a reliable alternative to the centralized schemes, a comparison between both schemes is presented.

Figure 8.8 presents the consumption profile of the part of the fleet considering the DR program of this work in both scenarios.

In the decentralized model, the aggregator has perfect information about the fleet requirements at every time step, and thus it reduces the cost of purchasing energy of the fleet optimally. As it can be observed, the load profile under a centralized approach is similar to the profile obtained with the scheme developed in this work, validating the decentralized model as a good alternative to the centralized model on buying energy.

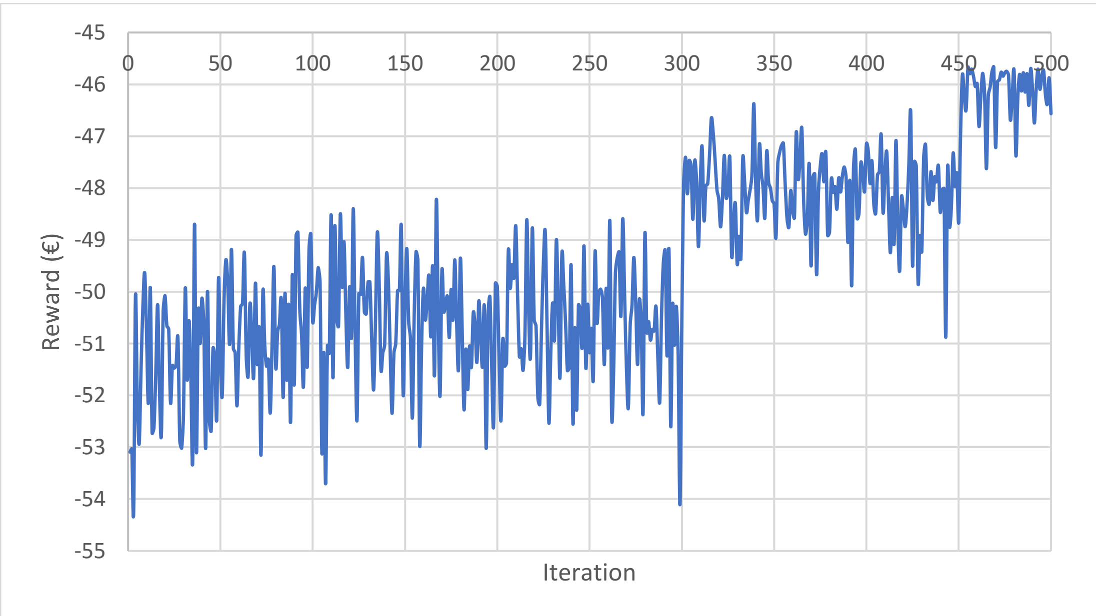

Figure 4.7 - Evolution of the reward of the 30 % of the fleet considering DR

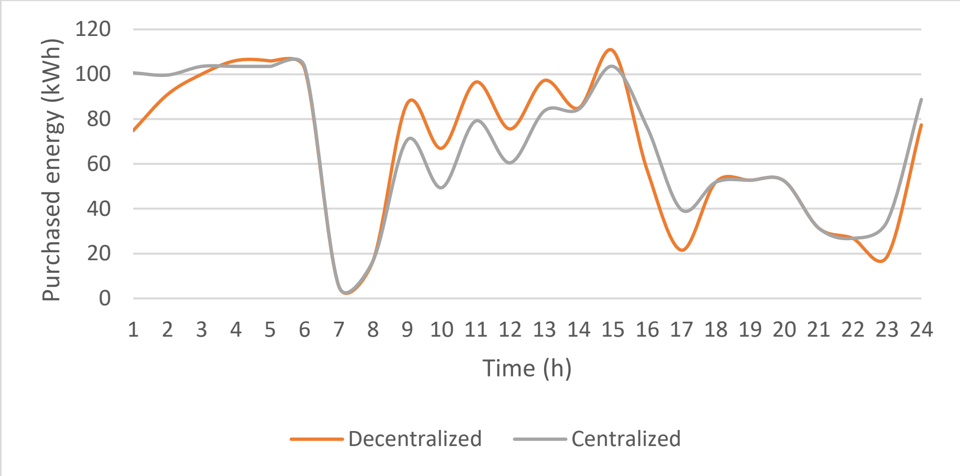

Figure 4.8 - Purchased energy, 30 % DR penetration

Figures 4.9 and 4.10 demonstrate that the decentralized scheme of this work leads to a considerably small deviation between the market clearing price and traded volumes throughout the day.

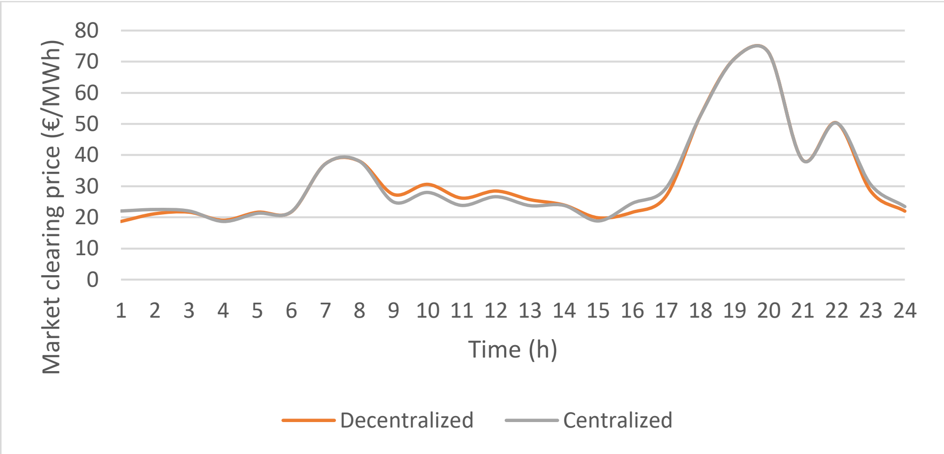

Figure 4.9 - Market clearing price, 30 % DR penetration

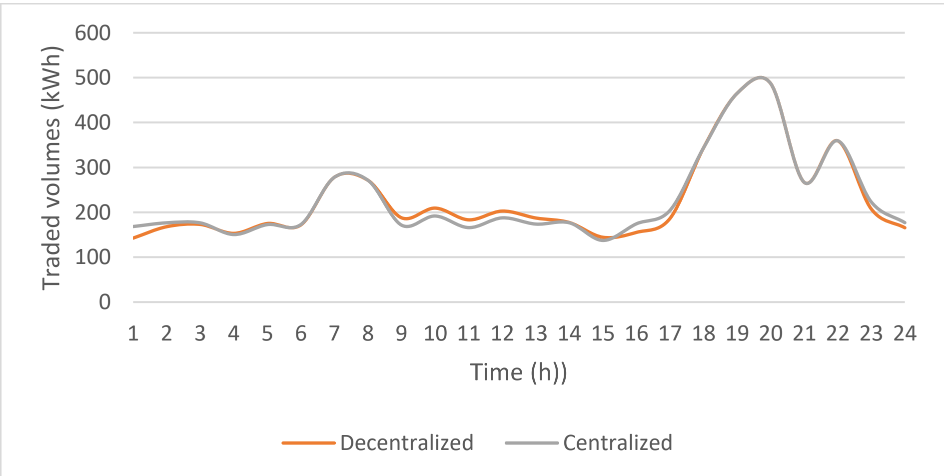

Figure 4.10 - Traded volumes, 30 % DR penetration

Table 4.4 presents the costs of purchasing energy per unit for the fleet considering DR. It shows that the results obtained with the decentralized scheme do not differ considerably from the results obtained with a centralized scheme.

Table 4.4 - Average cost of purchasing energy, decentralized vs. centralized model, in €/MWh

Decentralized model	Centralized model
29,75	28,10

4.2. 15% DR penetration

In the second DR scenario to be tested, a level of 15% DR penetration is considered. This value represents a scenario that is realistic in a near future, in which the possibility to participate in DR programs by the end-users has risen enough to lead the market to be influenced by the considered level.

- **Market Clearing Price:**

Figure 4.11 presents the market clearing price for the scenario in which all the loads of the fleet are inflexible and the scenario in which 15 % of the agents participate in the DR program developed in this work. As expected, the introduction of 15 % to the DR program of this work led to a reduction in the market clearing price during peak hours and to a raise during valley hours.

The difference between maximum and minimum clearing price is reduced by 18.01 % by using the DR model of this work, when comparing to the inflexible demand approach, as it can be observable in Table 4.5.

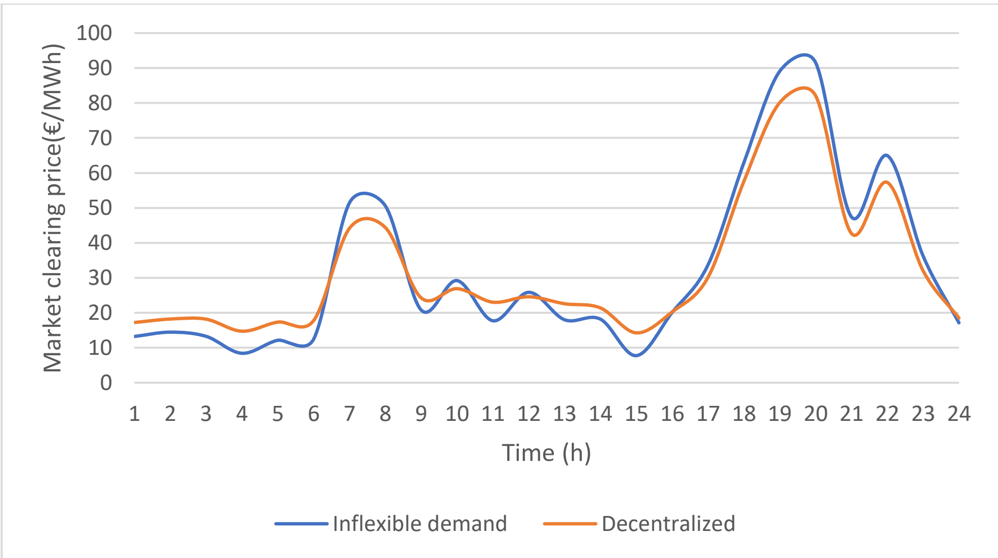

Figure 4.11 - Market clearing price, 15 % DR penetration

Table 4.5 - Variation between maximum and minimum MCP, 15 % DR penetration, in €/MWh

	Inflex. Demand	15 % DR
Min	7,74	13,38
Max	91,60	82,14
Max -Min	83,86	68,76
Variation (%)	18,01%	

- **Consumption profiles:**

Figure 4.12 shows the purchased energy of the 15 % of the fleet considered in this analysis, without demand response and considering the decentralized DR scheme of this work.

In terms of variation between the maximum and minimum value of purchased energy throughout the day, using the DR model of this work leads to a reduction of the variance between maximum and minimum amount of purchased energy by 21.30 % when compared to the uncontrolled approach, as it can be seen in Table 4.6.

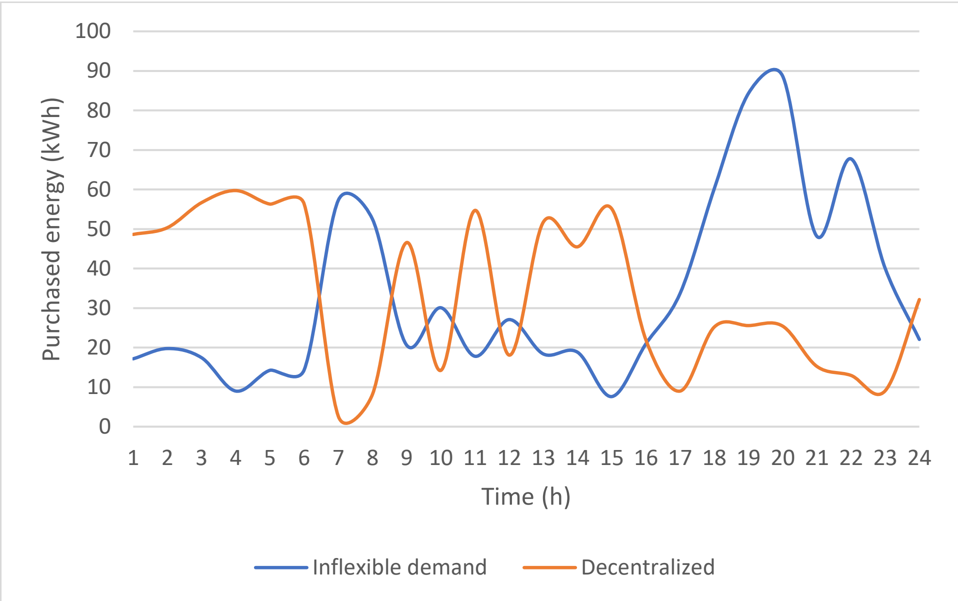

Figure 4.12 - Purchased energy, 15 % DR penetration

Table 4.6 - Variation between maximum and minimum quantity of bought energy, 15 % DR penetration, in kWh

	Inflex. Demand	15 % DR
Min	7,61	2,65
Max	88,70	66,47
Max - Min	81,09	63,82
Variation (%)	21,30%	

- **Traded volumes:**

Figure 4.13 presents the traded volumes for the scenario without demand response and for the scenario considering 15 % of the fleet to be participating in the decentralized DR program of this work. Regarding Table 4.7, it is noticeable that introducing 15 % of the fleet to the DR program of this work will lead to a reduction of the difference between maximum and minimum traded volumes of 18.78 % when compared to the inflexible demand approach.

By reducing the percentage of the fleet participating in the DR program to half, the variation between the difference of the minimum and maximum value for traded volumes with and without DR reduces from 37.95 % to 18.78 %. Thus, decreasing the percentage of DR participation of the fleet leads to a lower variation between the traded volumes throughout the day.

- **Costs of purchasing energy:**

Once more, the average cost of the purchased energy by the agents is calculated throughout the formulation in (21). In Figure 4.14 it is possible to observe the average cost of purchasing energy throughout the iterative process for the part of the fleet participating in the DR program of this work.

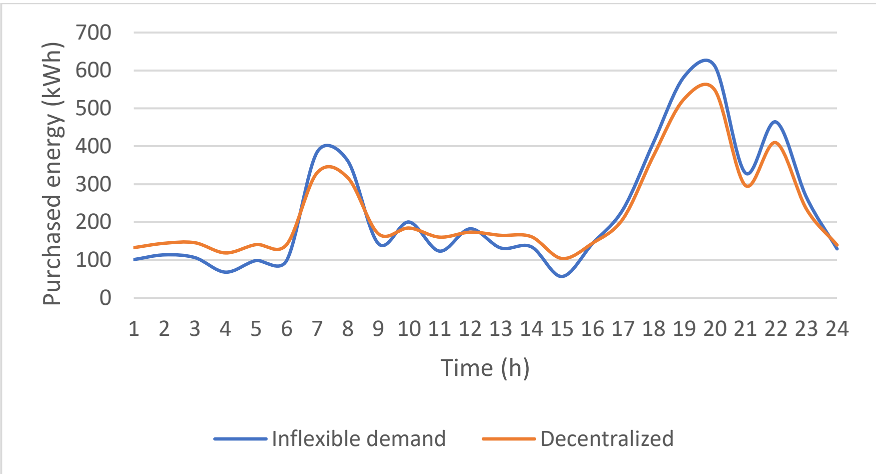

Figure 4.13 - Traded volumes, 15 % DR penetration

Table 4.7 - Variation between maximum and minimum traded volumes, 15 % DR penetration, in kWh

	Inflex. Demand	15 % DR
Min	56,34	97,37
Max	611,60	548,34
Max - Min	555,26	450,98
Variation (%)	18,78%	

The average cost for buying energy without demand response is 48.67 €/MWh. The average cost for purchasing energy by using the model developed in this work diminishes throughout the iterative process, as agents start defining the optimal actions, leading to a final average cost around 27.47 €/MWh. Thus, in a scenario with 15 % DR penetration the model developed in this work demonstrates to be considerably effective on diminishing the costs of purchasing energy.

- **Error:**

In Figure 4.15 it is possible to observe the difference between E_{exp} and E_{24} for each agent of the fleet. As it can be observable in Figure 4.15, the variance between the expected value of energy to purchase and the real value is, once more, not severely significant.

Figure 4.16 shows the variation between E_{exp} and E_{24} in terms of percentage. As it can be seen in Figure 4.16, during some hours of the day, the variation between both values is considerable, such as in hour 9. However, the variation between E_{exp} and E_{24} is also very low, such as in the last subsection, proving the efficiency of the model to obtain a small deviation between E_{exp} and E_{24} in a scenario with 15 % DR penetration.

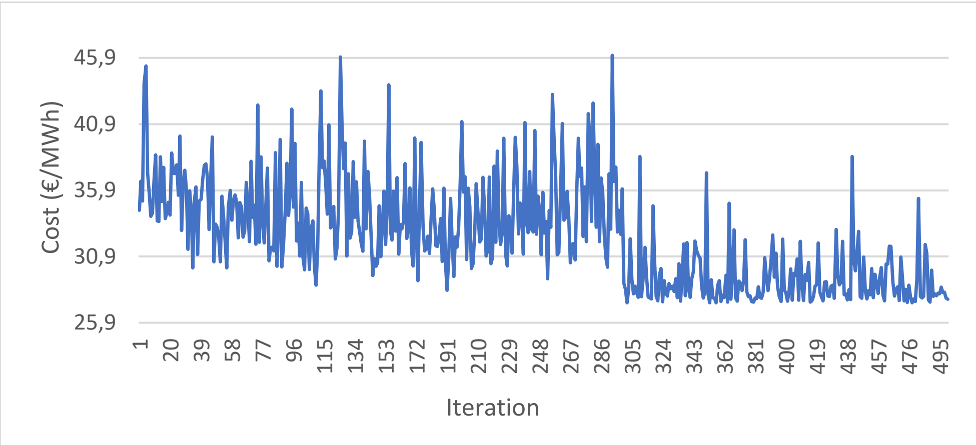

Figure 4.14 - Average cost of purchasing energy

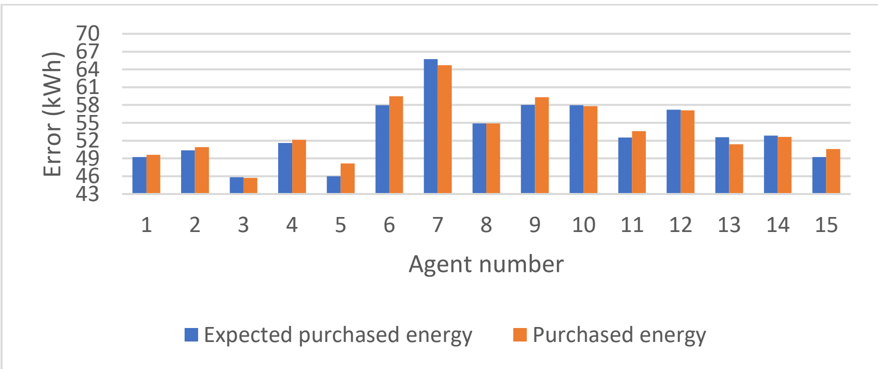

Figure 4.15 - Difference between expected purchased energy and the real value of purchased energy

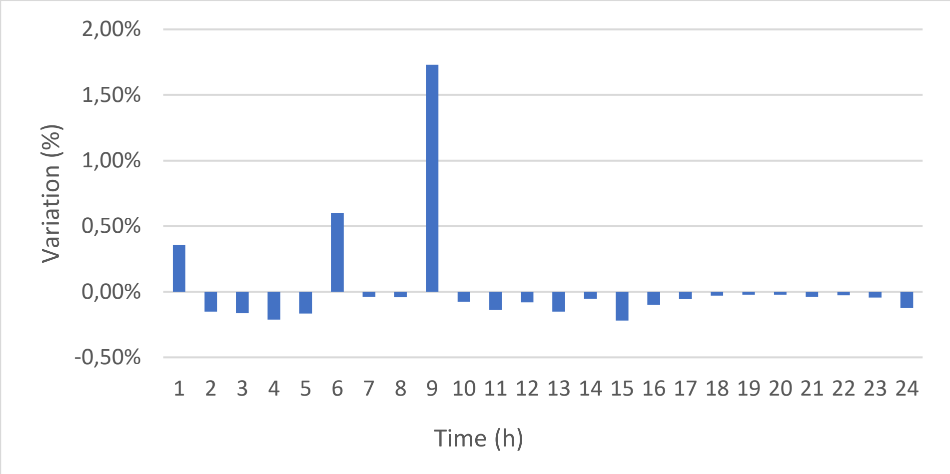

Figure 4.16 - Variation between assigned quantity of energy and purchased quantity of energy

- **Evolution of the reward:**

Figure 4.17 demonstrates the evolution of the fleet's reward throughout the iterative process. Once more, it is possible to observe that after a certain number of iterations, the value of the reward tends to converge to its optimal value, as Figure 4.17 demonstrates.

4.2.1. Centralized vs. Decentralized scheme

Figure 4.18 presents the energy purchased throughout the day by the 15 % of the fleet considering DR, both using the decentralized approach and the centralized one. Similarly to Figure 4.8, the load profile of the 15 % of the fleet using both approaches is similar, especially during peak hours. During valley hours, the variance is more noticeable because it is during those hours agents tend to buy energy to satisfy their shiftable loads, and therefore schedule their load consumption according to the followed approach.

Figure 4.19 and Figure 4.20 demonstrate that under a different level of DR penetration, both market clearing price and traded volumes barely vary from one approach to the other. Both values for the market clearing price and the traded volumes barely vary throughout the day, similarly to the scenario with 30 % DR penetration.

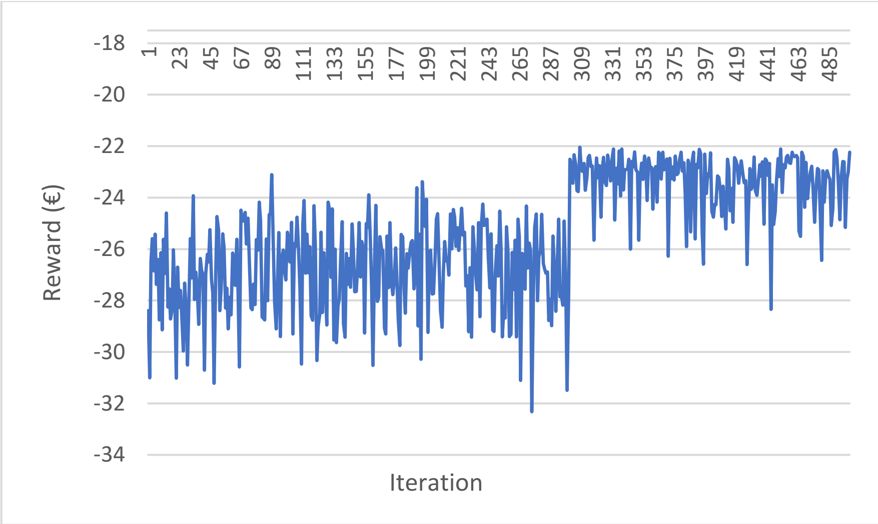

Figure 4.17 - Reward of the 15 % of the fleet considering DR

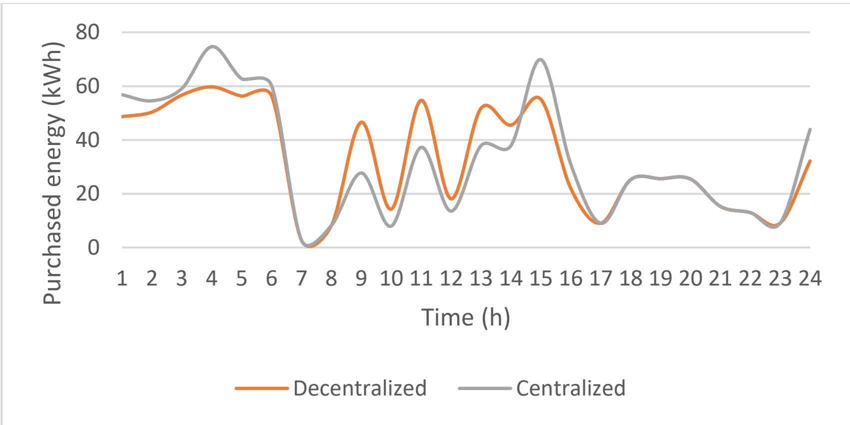

Figure 4.18 - Purchased energy, 15 % DR penetration

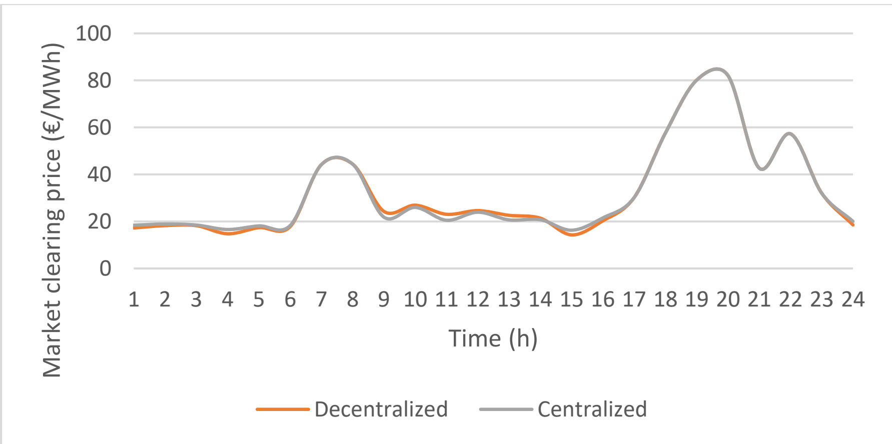

Figure 4.19 - Market clearing price, 15% DR penetration

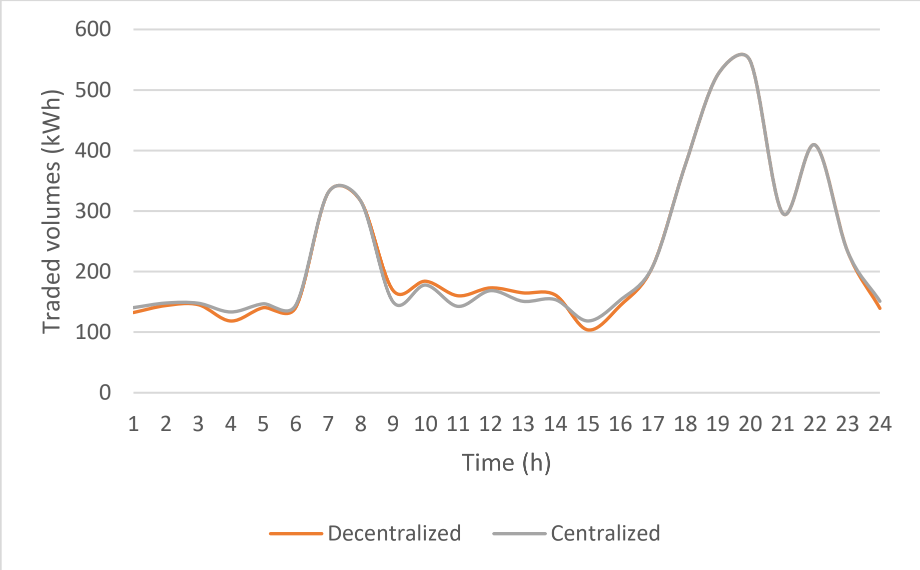

Figure 4.20 - Traded volumes, 15% DR penetration

Table 4.8 presents the costs of purchasing energy per unit for the fleet considering DR. It shows that the results obtained with the decentralized scheme do not differ considerably from the results obtained with a centralized scheme.

4.3. 60% DR penetration

In both scenarios studied in the previous chapters, the part of the fleet considering the demand to be inflexible dominated over the part of the fleet considering DR. Both scenarios are more realistic in a near future, in which DR is still considered as a novelty. Therefore, it is highly important to observe the behavior of this model in a scenario considering DR to be main approach to follow when it is about buying energy.

This scenario considers 60 % of the fleet to be participating in the DR program developed in this thesis, while the other 40 % remain as inflexible loads.

- **Market clearing price:**

Figure 4.21 presents the market clearing price throughout the day both without DR and considering the model developed in this work.

Table 4.8 - Average cost of purchasing energy, decentralized vs. centralized model, 15% DR penetration, in €/MWh

Decentralized model	Centralized model
27,47	25,74

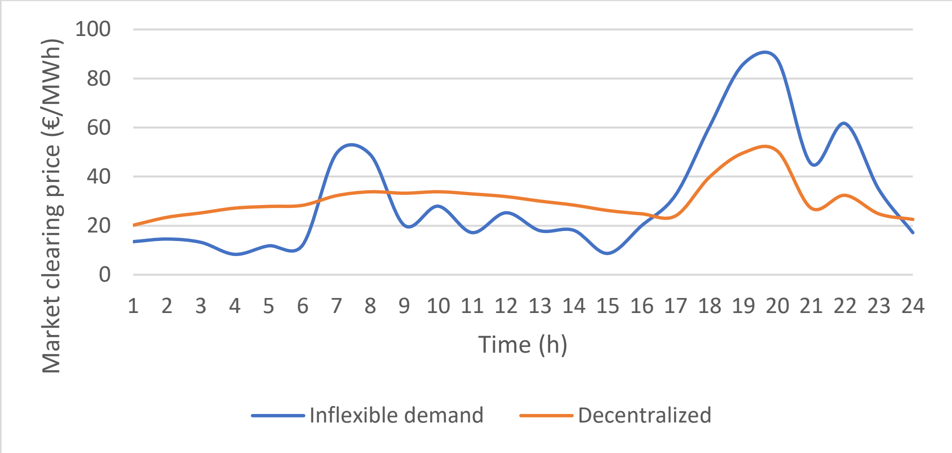

Figure 4.21 - Market clearing price, 60% DR penetration

In the scenarios studied before considering the DR level to be 60%, variations between the minimum and the maximum clearing price of the market were 18.78% and 37.95 %, for the DR levels of 15% and 30% respectively, when comparing to an approach without DR. Even considering the variation of these two levels to be considerably low when compared to an approach without demand response, whose variation between maximum and minimum is 48.67 €/MWh, it is possible to observe in Figure 4.21 that the price curve throughout the day is more balanced, with a variation between minimum and maximum value of 61.87 % when comparing to an approach without DR. This happens due to the fact that DR dominates over inflexible demand, leading the majority of the fleet to purchase energy evenly throughout the day, as for more than half of the fleet has not only sensitivity to the price to pay for energy during the day but also influence in the market clearing price.

Table 4.9 presents the variation between maximum and minimum clearing price in a scenario with 60 % DR penetration level and in a scenario with inflexible demand. As in can be observable, introducing 60 % of the fleet to the DR program of this work led to a considerable reduction in the variation between maximum and minimum market clearing price.

- **Consumption profile:**

Figure 4.22 shows the purchased energy of the 60 % of the fleet considered in this analysis, without demand response and considering the decentralized DR scheme of this work.

Table 4.9 - Variation between maximum and minimum MCP, 60 % DR penetration, in €/MWh

	Inflex. Demand	60 % DR
Min	8,31	20,18
Max	87,74	50,46
Max - Min	79,42	30,28
Variation (%)	61,87%	

In terms of variation between the maximum and minimum value of purchased energy throughout the day, using the DR model of this work leads to a reduction of the variance between maximum and minimum amount of purchased energy by 54.59 % when compared to the uncontrolled approach, as it can be seen in Table 4.10.

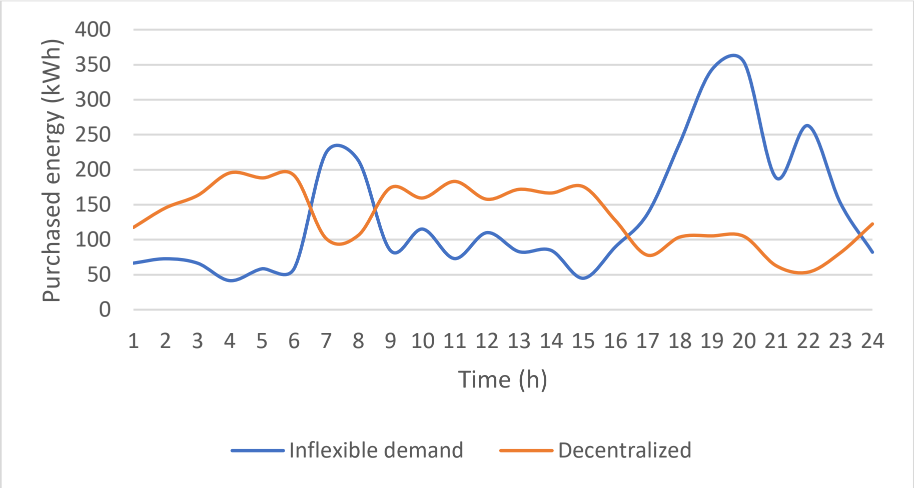

Figure 4.22 - Purchased energy, 60% DR penetration

Table 4.10 - Variation between maximum and minimum quantity of bought energy, 60 % DR penetration, in kWh

	Inflex. Demand	60 % DR
Min	41,47	53,59
Max	353,73	195,40
Max - Min	312,25	141,81
Variation (%)	54,59%	

- **Traded volumes:**

Figure 4.23 shows the traded volumes of energy throughout the day for both studied scenarios.

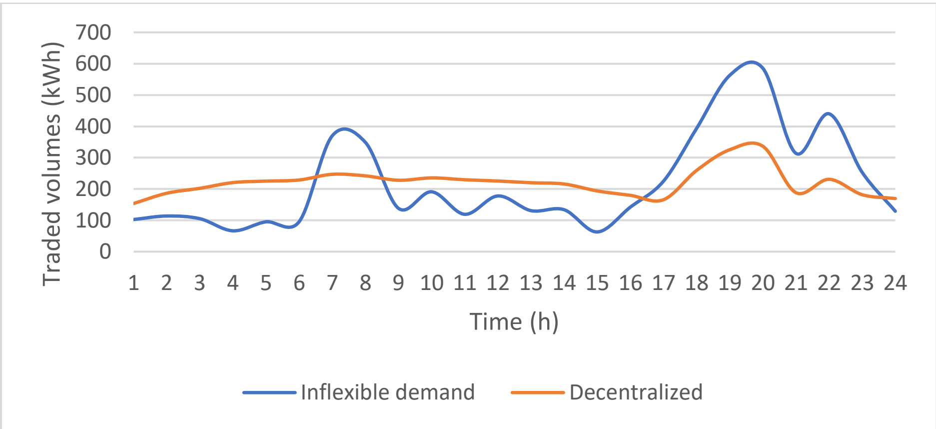

Figure 4.23 - Traded volumes, 60% DR penetration

Following the market clearing price variation, the difference of the minimum and maximum traded volume varied accordingly, as the price is dictated by the amount of purchased energy by the entire fleet.

Table 4.11 presents the variation between maximum and minimum traded volume of energy in both approaches. A variation of 65.00 % is explained by the fact that the number of end-users participating in the DR program is superior to the number of users whose demand is inflexible, and thus DR dominates over inflexible loads, leading to a huge impact in which concerns to leading the load diagram to be more even throughout the day.

- **Costs of purchasing energy:**

Once more, the average cost of the purchased energy by the agents is calculated throughout the equation (21). In Figure 4.24 it is possible to observe the average cost of purchasing energy throughout the iterative process.

The average cost for buying energy without demand response is 46.54 €/MWh. The average cost by using the model developed in this work diminishes throughout the iterative process, as agents start defining the optimal actions, leading to a final average cost around 31.57 €/MWh. Thus, in a scenario with 60 % DR penetration the model developed in this work demonstrates to be considerably effective on diminishing the costs of purchasing energy.

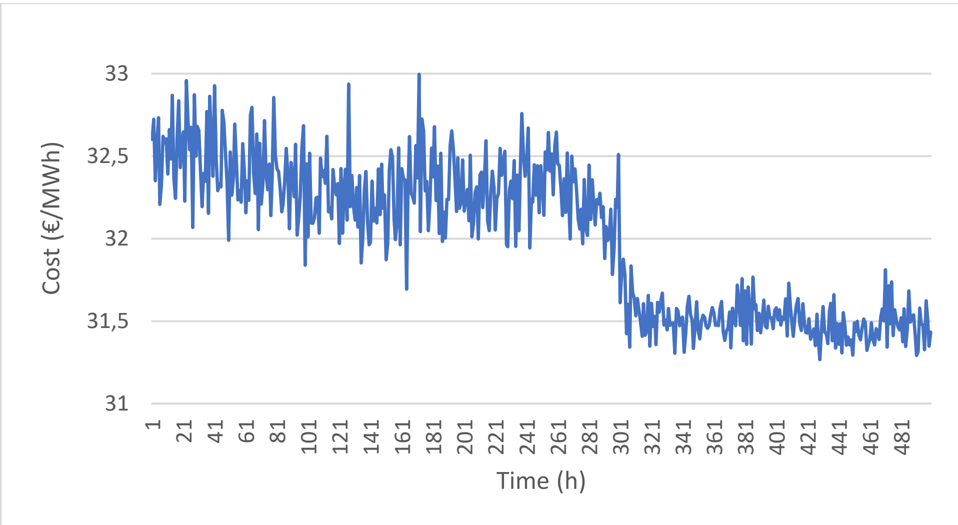

Figure 4.24 - Cost of purchasing energy

Table 4.11 - Variation between maximum and minimum traded volumes, 60 % DR penetration, in kWh

	Inflex. Demand	60 % DR
Min	62,95	153,64
Max	585,50	336,53
Max - Min	522,55	182,88
Variation (%)	65,00%	

- **Error:**

In Figure 4.25 it is possible to observe the difference between E_{exp} and E_{24} for each agent of the fleet. As it can be observable, the variance between the expected value of energy to purchase and the real value is, once more, not severely significant.

Figure 4.26 shows the variation between E_{exp} and E_{24} in terms of percentage. As it can be seen, during some hours of the day, the variation between both values is considerable, such as in hours 13 and 14. However, the variation between E_{exp} and E_{24} is also very low, such as in the last subsections, proving the efficiency of the model to obtain a small deviation between E_{exp} and E_{24} in a scenario with 60 % DR penetration.

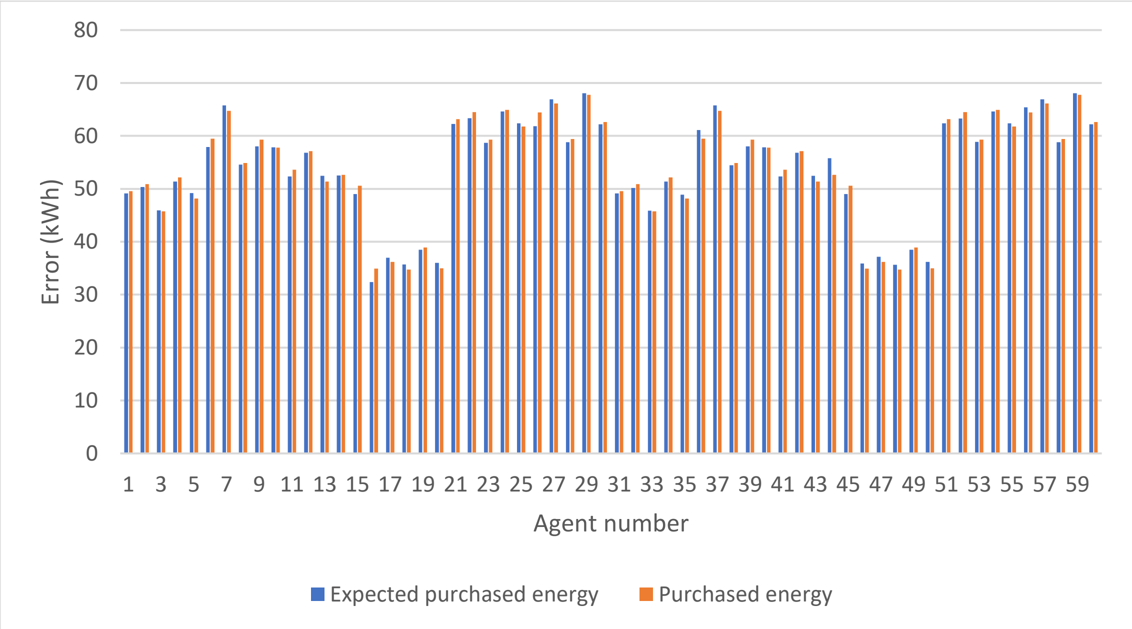

Figure 4.25 - Difference between expected purchased energy and the real value of purchased energy

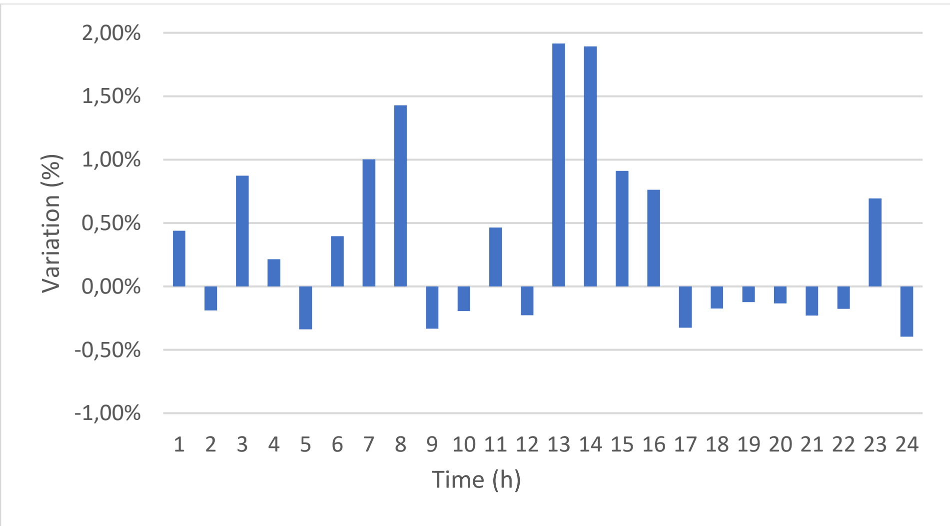

Figure 4.26 - Variation between assigned quantity of energy and purchased quantity of energy

- **Evolution of the reward:**

Figure 4.27 demonstrates the evolution of the reward during the iterative process. Once more, agents try a varied set of actions a sufficient number of times before starting to determine which actions are more profitable. After a certain amount of iterations, agents start performing the best actions, leading the results to be optimal in terms of diminishing the cost of purchasing energy of the fleet.

4.3.1. Centralized vs. Decentralized scheme

Figure 4.28 shows the purchased energy during the day of the 60 % of the fleet considering both DR schemes.

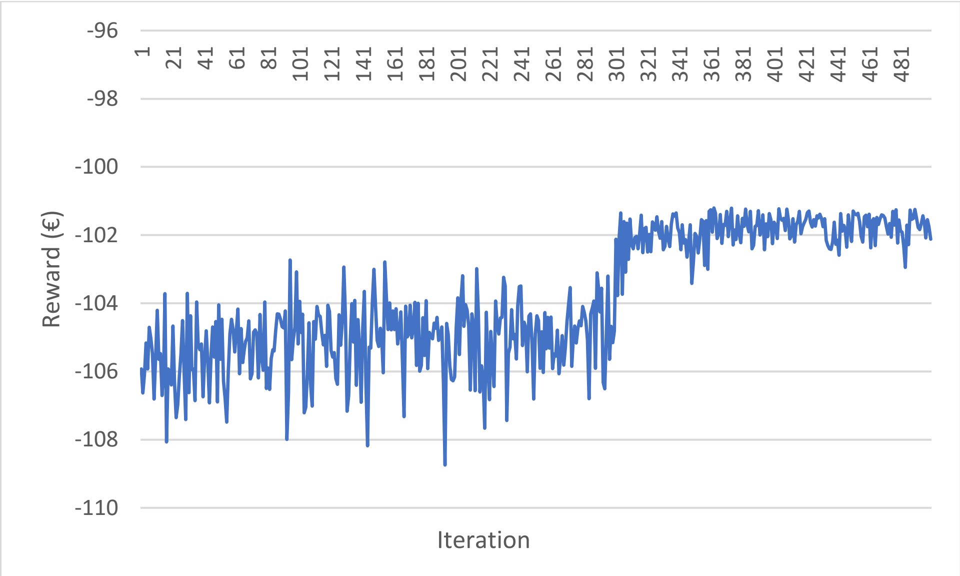

Figure 4.27 - Reward of the 60% of the fleet considering DR

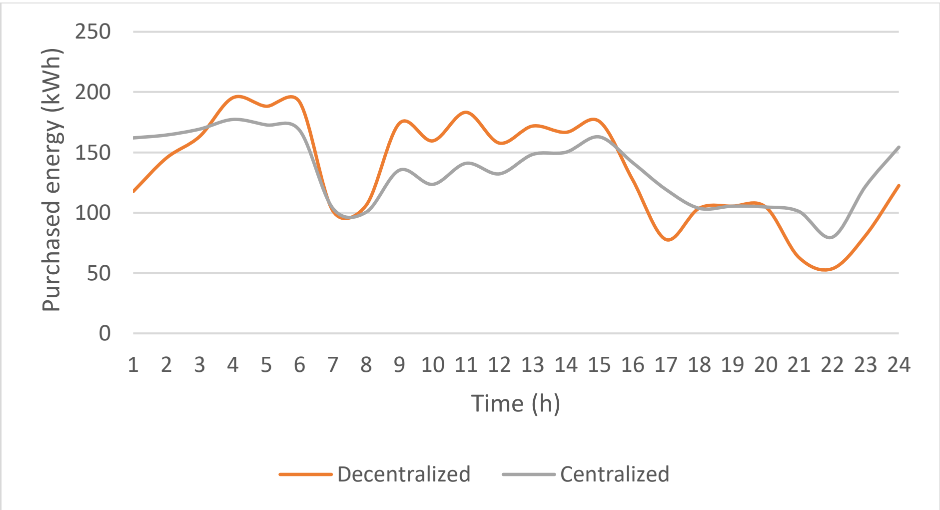

Figure 4.28 - Purchased energy, 60% DR penetration

Regarding Figure 4.28, the variation of purchased energy outside of peak hours is more noticeable than in the previous scenarios. This happens due to the fact than while the aggregator purchasing energy on behalf of the fleet in a centralized model has perfect information. On the other side, increasing the size of the DR fleet increases the uncertainty in which concerns to price to pay for energy, as there are more agents influencing the market clearing throughout their individual bids, and therefore more possibilities in which concerns to the amount of energy to buy each hour.

That same size variation also reflects on the market clearing price and traded volumes outside of valley hours, as it can be observable in Figures 4.29 and 4.30.

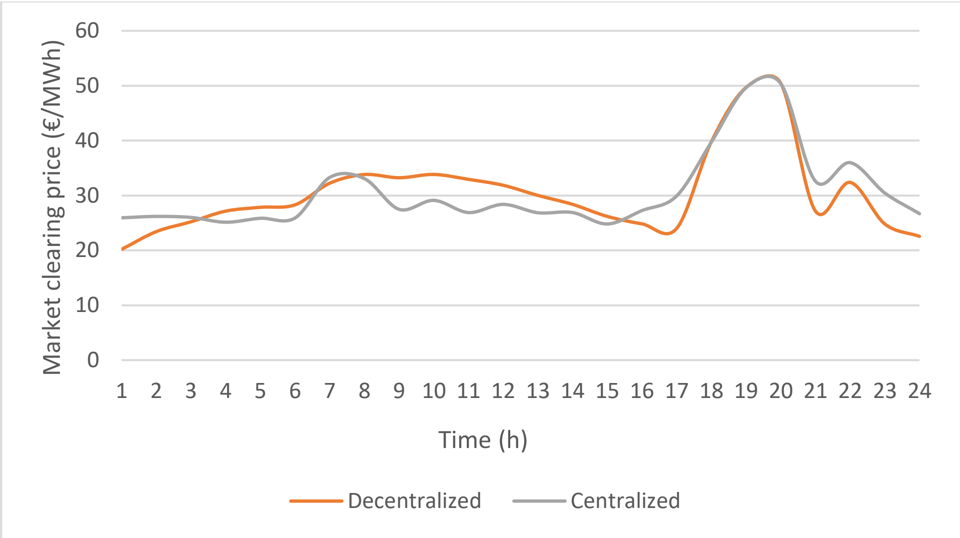

Figure 4.29 - Market clearing price, 60% DR penetration

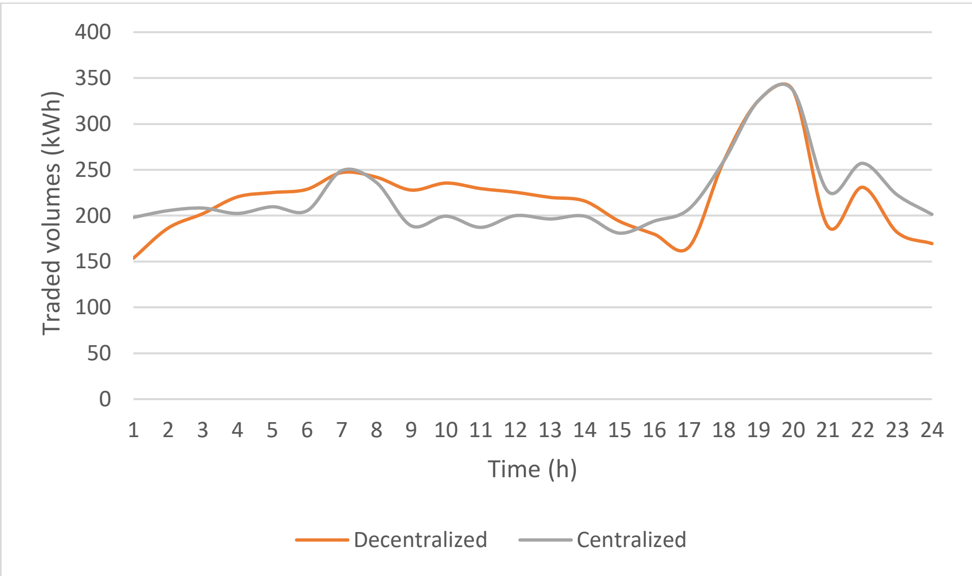

Figure 4.30 - Traded volumes, 60% DR penetration

Table 4.12 presents the costs of purchasing energy per unit for the fleet considering DR. It shows that the results obtained with the decentralized scheme do not differ considerably from the results obtained with a centralized scheme.

4.4. Results comparison

The main goal of this work is to model a decentralized market-based scheme for DR, being the main goal to diminish the costs of purchasing energy. However, another problem associated with inflexible demand is the high demand for energy during peak hours, leading the network to considerable problems in which concerns to provide energy to the buyers in situations such as the failure of a generator. Correctly dimensioned DR programs lead the demand during peak hours to diminish, being easier for the network to provide end-users energy when there is a risk of demand to surpass the maximum level of energy the network can provide. Thus, another focus of this work is to study the impact of the decentralized model in the general load diagram considering the penetration level of DR.

Figure 4.31 demonstrates the average purchased energy each hour by each agent. As it can be observable, the higher the penetration level of DR, the lower the variation of purchased energy during the day. To test the efficiency of the model developed in this work, the total cost of buying energy throughout the usage of the model is compared to an approach without demand response, in which all the loads are defined as inflexible, and a centralized approach, in which the aggregator has perfect information, and therefore is capable to reduce the costs of buying energy to its optimal value.

Table 4.12 - Average cost of purchasing energy, decentralized vs. centralized model, 60% DR penetration, in €/MWh

Decentralized model	Centralized model
31,57	29,58

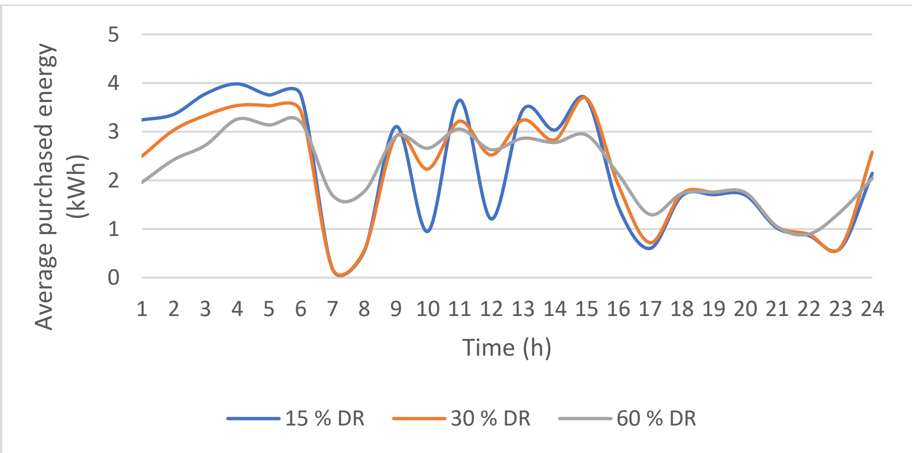

Figure 4.31 - Average purchased energy per agent

Table 4.13 shows the total cost of buying energy of the fleet within the three different DR penetration levels and the three different approaches. As it can be observable, the decentralized model is effective independently from the DR penetration level. While the decentralized model diminishes the cost of purchasing energy considerably when comparing to an uncontrolled approach, its values for the cost of buying energy don't differ substantially from the centralized approach values.

Thus, it can be concluded that the decentralized model developed in this work is a very reliable alternative to the centralized model in which concerns diminishing the costs of purchasing energy, while solving the issues related to the centralized approach such as the privacy of the agents.

Table 4.14 presents the average cost of purchasing energy for the part of the fleet considering DR, regarding all three scenarios of DR penetration. As it can be observable, the average cost obtained throughout the usage of the model developed in this work diminishes the average cost considerably in comparison to the scenario with inflexible demand, and the deviation between both centralized and decentralized models is not accentuated.

As the DR penetration level raises, the average cost of purchasing energy slightly raises as well. This is due to the fact that the bigger the number of participants in the DR programs, the bigger the competition for energy is during the hours DR agents purchase energy. However, increasing the number of participants doesn't affect the average cost of energy linearly, as increasing the number of participants from 15 % to 60 % only increases the cost by 12.97 % for the decentralized model, which leads to the conclusion that as the number of participants in the DR program grows the price to pay for energy grows as well, but not substantially enough as to lead the model to become obsolete.

Thus, it can be concluded that not only the new decentralized model of this work not only reduces the cost of purchasing energy close to the optimality, but also that increasing the percentage of participants will not considerably affect the cost reduction when comparing to an approach without DR.

It also helps diminishing the variation of minimum and maximum amount of purchased energy throughout the day, helping the network operators on providing energy in case of failure on serving loads during peak hours, and dealing with the growth of the energy fleet that leads to a raise in the demand for energy.

Table 4.13 - Total cost of purchasing energy in all three approaches and all three DR level, in €

DR level	Inflexible demand	Decentralized approach	Centralized approach
15%	39,32 €	22,03 €	20,80 €
30%	78,44 €	48,23 €	45,58 €
60%	150,97 €	102,20 €	95,94 €

Table 4.14 - Average cost of purchasing energy in all three approaches and all three DR level for the fleet considering DR, in €/MWh

DR level	Inflexible demand	Decentralized model	Centralized model
15%	48,67 €/MWh	27,47 €/MWh	25,74 €/MWh
30%	48,37 €/MWh	29,75 €/MWh	28,10 €/MWh
60%	46,54 €/MWh	31,57 €/MWh	29,58 €/MWh

Chapter 5

Conclusion

A new decentralized market-based scheme considering demand response was developed in the present thesis, being the results obtained with the model compared to the ones obtained in an approach without demand response, and the ones obtained with a centralized model.

The main goal of the thesis was to develop a bidding mechanism in which each agent defines its bids according to their urgency for buying energy, being the bids optimally placed throughout the usage of a Q-learning algorithm. Agents buy energy according to the market clearing price and their previously defined demand bid curves.

In which concerns to the Q-learning algorithm, agents demonstrate to considerably optimize their bids, as the algorithm converges to optimal values as the iterative process follows through. To test the efficiency of the model in various near future scenarios, three levels of DR penetration were considered. For the three considered levels, the model demonstrates to be effective on reducing the costs of purchasing energy when compared to an approach without demand response. Agents participating in the DR program of this work buy more energy during hours in which the market clearing price is lower, diminishing their consumption during peak hours.

Regardless the DR penetration level, the developed model not only reduces the cost of energy considerably, but also diminishes the variation between the maximum and minimum value for traded volumes of energy. As the DR level ascends, the same variation demonstrates to diminish, proving the efficiency of the method on leading the load diagram of the entire fleet to be more balanced, reducing problems related to high levels of demand during peak hours. When comparing the results to the ones obtained with a centralized model, it is observable that the results are considerably approximate and slightly higher. When comparing the market clearing price profiles and the load consumptions profiles of both methods, the variation is small.

Thus, it can be concluded that the model developed in this work not only reduces substantially the costs of purchasing energy in a scenario without DR, but also is a reliable alternative to the centralized models, solving the scalability and privacy problems related with this type of model. The new decentralized model of this work demonstrates to be an approach closer to the one to be developed in a near future, in which a bigger amount of household consumers will have access to DR programs.

5.1. Future work

The concept of DR is a fresh topic regarding the fact that few people have access to programs at the present time. Previous DR models either have problems related to scalability or load synchronization. Thus, the innovation to me made in this field has a wide range of options to be explored and developed in a near future.

Regarding the work developed in this thesis, one interesting feature would be considering consumption uncertainty, as to test the impact in the cost of purchasing energy in such scenario. To approach the model to real market situation, both uncertainty in the demand and supply could be considered, using real market data so the obtained results would be closer to the ones to be observed in the future. As mentioned in [42-46], many algorithms were proposed as a way to determine the energy prices and the costumers' responses to them. In these scenarios, supply and demand are considered to match DR being the aggregated demand equalized to the supply. However, since households aren't isolated but in fact connected to energy networks, another interesting feature would be integrating the model into an OPF market clearing, similarly to [47].

Considering the model of this work is directed to household consumers, an algorithm as the one presented in [48] for intelligent home energy management could be introduced so end-users could define their load serving priorities. Throughout the usage of the algorithm, household appliances would be served in a particular order previously defined by the users, and load serving would be done not only according to their individual demand bids and the market price but also the load priority and also the particular urgency to serve a particular load, i.e., the state of every individual load would be considered, as for the agents to buy energy according to their comfort preference in a more detailed way.

Another interesting development would be adapting the three-block demand bid curves to a more detailed one, in which every user sets, as an example, one block for each load, according their priority, and serve every particular load according to the price, leading the agents to have not only a bigger degree of freedom, but also to purchase the exact amount of energy to serve a certain amount of loads, as opposition to the approach of this thesis, in which all the shiftable loads are simplified to a single one. As to model each load individually and in a more appropriate futuristically way, each load could be shaped as a physical-based load such as in [49-53].

References

[1] http://205.254.135.24/forecasts/ieo/pdf/0484(2011).pdf;

[2] Berger, L. T., and Iniewski, K., "*Smart Grid: Applications, Communications, Security*", Hoboken, NJ, USA: Wiley, Apr. 2012;

[3] "Assessment of demand response and advanced metering", Washington, DC, USA, Tech. Rep., Dec. 2012;

[4] Li, Y., Ng, B. L., Trayer, M., and Liu, L. "Automated residential demand response: Algorithmic implications of pricing models," *IEEE Trans. Smart Grid*, vol. 3, no. 4, Dec. 2012, p. 1712–1721;

[5] Ullah, I., Javaid, N., Imran, M., Khan, Z. A., Qasim, U., Alnuem, M., and Bashir, M., "A Survey of Home Energy Management for Residential Customers", *IEEE 29th International Conference on Advanced Information Networking and Applications*, 2015, p. 666-673;

[6] Aalami, H., Moghadam, M. P and Yousefi, G. R., "Optimum time of use program proposal for Iranian power systems", *Proc. International Conference on Electrical Power Energy Conversion Systems*, November 2009;

[7] Ashok, S., and Banerjee, R., "Optimal operation of industrial cogeneration for load management", *IEEE Transactions Power System*, vol. 18, no. 2, May 2003, p. 931-937;

[8] Joo, J., Ahn, S., Yoon, Y., and Choi, J., "Option valuation applied to implementing demand response via critical peak Pricing", *Proc. IEEE Power Energy Society General Meeting*, Jun. 2007;

[9] Philpott, A. B., and Pettersen, E., "Optimizing demand-side bids in day-ahead electricity market", *IEEE Trans. Power Syst.*, vol. 21, no. 2, May 2006, p. 488-498;

[10] http://www.ferc.gov/legal/staff-reports/2010-dr-report.pdf;

[11] Vayá, M. G., Roselló, L. B., and Andersson, G., "Optimal Bidding of Plug-In Electric Vehicles in a Market-Based Control Setup", *Power Systems Computation Conference*, 2014;

[12] Bessa, R. J., Matos, M. A., Soares, F. J., and Lopes, J. A. P., "Optimized bidding of a EV aggregation agent in the electricity market", *IEEE Transactions on Smart Grid*, vol. 3, no. 1, 2012, p. 443-452;

[13] Kristoffersen, T. K., Capion, K., and Meibom, P., "Optimal charging of electric drive vehicles in a market environment", *Applied Energy*, vol. 88, no. 5, 2011, p. 1940-1948;

[14] Vagropoulos, S. I., and Bakirtzis, A. G, "Optimal bidding strategy for electric vehicle aggregators in electricity markets," *IEEE Transactions on Power Systems*, vol. 28, no. 4, 2013, p. 4031–4041;

[15] Sundstrom, O., and Binding, C., "Flexible charging optimization for electric vehicles considering distribution grid constraints," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, 2012, p. 26–37;

[16] Vayá, M. G., and Andersson, G., "Optimal bidding strategy of a plug-in electric vehicle aggregator in day-ahead electricity markets", *International Conference on the European Energy Market*, 2013;

[17] Vayá, M. G., Roselló, L. B., and Andersson, G., "Centralized and decentralized approaches to smart charging of plug-in vehicles", *IEEE Power and Energy Society General Meeting*, 2012;

[18] Rotering, N., and Ilic, M., "Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets," *IEEE Transactions on Power Systems*, vol. 26, no. 3, 2011, p. 1021–1029;

[19] Bashash, S., Moura, S. J., Forman, J. C., and Fathy, H. K., "Plug-in hybrid electric vehicle charge pattern optimization for energy cost and battery longevity," *Journal of Power Sources*, vol. 196, no. 1, 2011, p. 541–549;

[20] A. Hoke, A. Brissette, D. Maksimovic, A. Pratt, and K. Smith, "Electric vehicle charge optimization including effects of lithium-ion battery degradation," in *IEEE Vehicle Power and Propulsion Conference, 2011*;

[21] Kok, J. K., Scheepers, M. J. J., and Kamphuis, I. G., "Intelligence in electricity networks for embedding renewables and distributed generations", *Intelligent Infrastructures, Intelligent Systems, Control and Automation: Science and Engineering,* vol. 42, Springer Netherlands, 2010, p. 179-209;

[22] Ma, Z., Callaway, D. S., and Hiskens, I. A., "Decentralized charging control of large populations of plug-in electric vehicles", *IEEE Transactions on Control Systems Technology,* vol. 21, no. 1, 2013, p. 67-78;

[23] Gan, L., Topcu, U., and Low, S., "Optimal decentralized protocol for electric vehicle charging", *IEEE Transactions on Power Systems*, vol. 28, no. 2, 2013, p. 940-951;

[24] Vandael, S., Claessens, B., Hommelberg, M., Holvoet, T., and Deconinck, G., "A scalable three-step approach for demand side management of plug-in hybrid vehicles", *Smart Grid, IEEE Transactions*, 4(2), 2013, p. 720-728;

[25] Shao, S., Pipattanasomporn, M., and Rahman, S; "Demand Response as a Load Shaping Tool in an Intelligent Grid with Electric Vehicles", *IEEE Transactions*, vol. 2, no. 4, December 2011, p. 624-631;

[26] Liu, Y., Holtzer, J. T., and Ferris, M. C., "Extending the bidding format to promote demand response", *Energy Policy 86*, 2015, p. 82-92;

[27] Watkins, C. J. C. H., "Learning from delayed rewards", PhD thesis, King's College, Cambridge, 1989;

[28] Lopez, M. A., Aguado, J. A., Torre, S., and Figueroa, M., "Optimization-Based Market-Clearing Procedure with EVs Aggregator Participation", *Innovative Smart Grid Technologies Europe, 4th IEEE/PES,* 2013;

[29] Krause, T., "Evaluating Congestion Management Schemes in Liberalized Electricity Markets Applying Agent-based Computational Economics", PhD thesis, Swiss Federal Institute of Technology, Zurich, 2007;

[30] MacArthur, S. D. J., Catterson, V. M., Hatziargyriow, N. D., and Funabashi, T., "Multi-Agent Systems for Power Engineering Applications – Part I: Concepts, Approaches, and Technical Challenges", *IEEE Transactions on Power Systems*, vol. 22, no. 4, November 2007, p. 1743-1752;

[31] Papavasiliou, A., Hindi, H., and Greene, D., "Market-based control mechanism for electric power demand response", *Decision and Control, 49th IEEE Conference*, 2010, p. 1891-1898;

[32] Siano, P., and Sarno, D., "Assessing the benefits of residential demand response in a real time distribution energy market", *Applied Energy 161*, 2016, p. 533-551;

[33] Haider, H. T., See, O. H., and Elmenreich, W.; "Residential demand response scheme based on adaptative consumption level pricing", *Energy 113*, 2016, p. 301-308;

[34] Bessa, R. J., and Matos, M. A., "Economic and technical management of an aggregation agent for electric vehicles: a literature survey", *European Transaction on Electrical Power*, 22(3), 2012, p. 334-350;

[35] Lopes, J. A. P., Soares, F. J., and Almeida, P. M. R., "Integration of electric vehicles in the electric power system", *Proceeding of the IEEE*, 99(1), 2011, p. 168-183;

[36] Vayá, M. G., and Andersson, G., "Locational marginal pricing based smart charging of plug-in hybrid vehicle fleets", *Smart Energy Strategies Conference*, 2011, p. 21-23;

[37] https://pt.khanacademy.org/economics-finance-domain/macroeconomics/aggregate-supply-demand-topic/aggregate-supply-demand-tut/v/aggregate-demand;

[38] Murphy, K. P., "Machine learning: a probabilistic perspective", *MIT Press*, 2012;

[39] Sutton, R. S., and Barto, A. G., "Reinforcement Learning: An Introduction", vol. 2, *MIT Press*, London, UK, 2012;

[40] https://www.khanacademy.org/economics-finance-domain/microeconomics/supply-demand-equilibrium/market-equilibrium-tutorial/v/market-equilibrium;

[41] Krause, T., Beck, E. V., Cherkaoui, R., Germond, A., Andersson, G., and Ernst, D., "A comparison of Nash equilibria analysis and agent-based modelling for power markets," *International Journal of Electrical Power & Energy Systems*, vol. 28, no. 9, 2006, p. 599-607;

[42] Samadi, P., Mohsenian-Rad, A. H., Schober, R., Wong, V. W. S., and Jatskevich, J. "Optimal real-time pricing algorithm based on utility maximization for smart grid,", *Proc. IEEE SmartGridComm*, Gaithersburg, MD, USA, Oct. 2010, p. 415–420;

[43] Chen, L., Li, N., Low, S. H., and Doyle, J. C., "Two market models for demand response in power networks,", *Proc. IEEE SmartGridComm*, Gaithersburg, MD, USA, Oct. 2010, pp. 397–402;

[44] Yu, R., Yang, W., and Rahardja, S., "A statistical demand-price model with its application in optimal real-time price," *IEEE Trans. Smart Grid*, vol. 3 no. 4, Dec. 2012, p. 1734–1742;

[45] Li, N., Chen, L., and Low, S. H., "Optimal demand response based on utility maximization in power networks," *Proc. IEEE PES Gen. Meet.*, Detroit, MI, USA, Jul. 2011, p. 1-8;

[46] Qian, L. P., Zhang, Y., Huang, J., and Wu, Y., “Demand response management via real-time electricity price control in smart grids,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, Jul. 2013, p. 1268–1280;
[47] Mohsenian-Rad, H., and Davoudi, A., “Optimal demand response in DC distribution networks,” in *Proc. IEEE SmartGridComm*, Vancouver, BC, Canada, Oct. 2013, p. 564–569;
[48] Shi, W., Li, N., Xie, X., Chu, C., and Gadh, R., “Optimal Residential Demand Response in Distribution Networks”, *IEEE Journal on Selected Areas in Communication*, vol. 32, no. 7, 2014, p. 1441-1449;
[49] Ihara, S., and Schweppe, F. C., “Physically based modeling of cold load pickup,” *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 9, Sep. 1981, p. 4142–4150;
[50] Agneholm, E., and J. Daalder, “Cold load pick-up of residential load,” *Proc. Inst. Elect. Eng., Gen., Transm., Distrib.*, vol. 147, no. 1, Jan. 2000, p. 44–50;
[51] Paull, L., MacKay, D., Li, H., and Chang, L., “A water heater model for increased power system efficiency,” in *Proc. Canadian Conf. Electrical and Computer Engineering*, May 3–6, 2009;
[52] Gomes, A., Antunes, C. H., and Martins, A. G., “Physically-based load demand models for assessing electric load control actions,” in *Proc. IEEE Bucharest PowerTech*, Jul. 2009;
[53] Mortensen, R. E., and Haggerty, K. P., “A stochastic computer model for heating and cooling loads,” *IEEE Trans. Power Syst.*, vol. 3, no. 3, Aug. 1988, p. 1213–1219.

Appendices

A.1 Actions

This appendix serves to demonstrate the influence actions may have on finding solutions close to the optimal ones, as to apply the most profitable actions during the simulations of the results presented in Chapter 4. All tests were performed merely to the 30 % DR penetration level.

A.1.1. Testing B_1

To test the influence of the parameter B_1 on reducing the costs of purchasing energy, two different scenarios are considered. In Scenario 1, B_1 has a set of nine values between 10 and 90, two below the market clearing prices without DR (4 and 5), and two above the same market clearing price range (100 and 105). Three values of B_2 are defined, and thus a total of 39 actions are set. In Scenario 2, the most popular actions regarding Scenario 1 are tested as to observe the influence in the cost of purchasing energy. Five different values are defined for B_1 between 10 and 30, and the same three values for B_2 are defined, leading to a settlement of 15 actions in total.

As in can be seen, considering merely the most profitable actions reduces the amount of purchased energy during the first hours of the day. In which concerns to the average cost of purchasing energy, Table A.1 presents the average cost for both scenarios for the 30 % of the fleet considering DR, leading to the conclusion, Figure A.1 presents the bought amount of energy for both scenarios. reducing the number of actions to a set of values close to the most popular actions reduces deviations between the expected amount of energy to purchase and the actual value for the purchased energy. Figure A.2 demonstrates the amount of purchased energy of one agent of the fleet. As it can be seen, when the set of actions is more adequate and approached to the most profitable actions, the agent tends to purchase energy more evenly throughout the day.

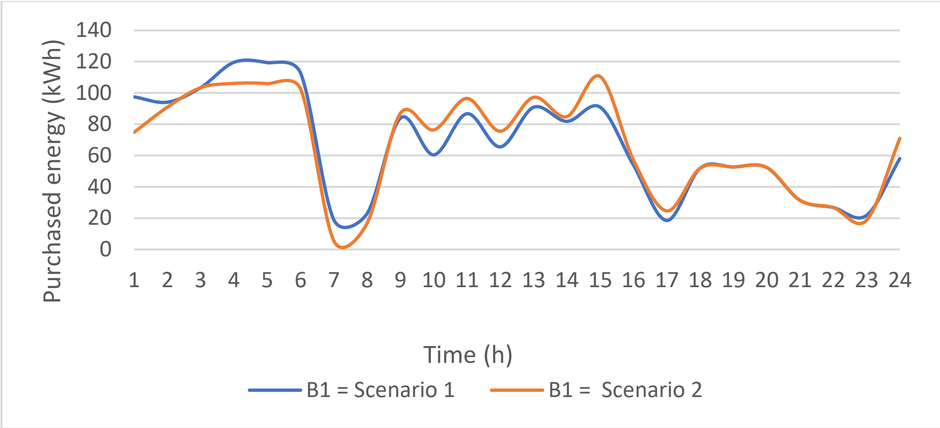

Figure A.1 - Purchased energy in both considered scenarios

Table A.1 - Average cost or purchasing energy for both considered scenarios, in €/MWh

Scenario 1	Scenario 2
30,96	29,75

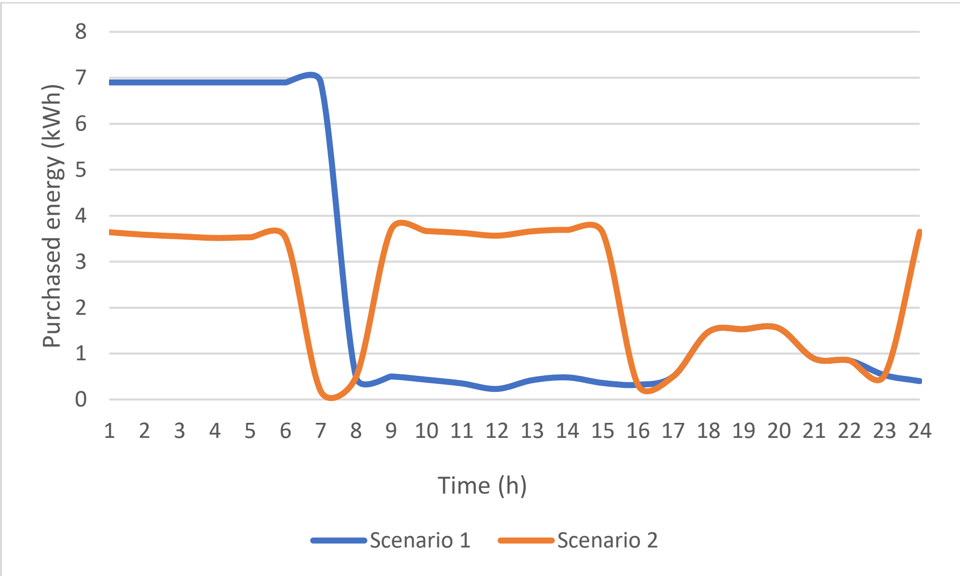

Figure A.2 - Purchased energy of one of the agents of the fleet considering both scenarios, in kWh

A.1.2. Testing B_2

As to test the influence of B_2 in the cost of purchasing energy, three values are defined, 6.9, 10.35, and 13.8. The first scenario only considers 10.35, being the two other values attributed to the second scenario as to test the impact of extending the range of actions.

Figure A.3 presents the purchased amount of energy for both scenarios. It is observable that for both scenarios the amount of purchased energy doesn’t differ considerably. Table A.2 presents the average cost of purchasing energy for both scenarios. A wider range for B_2 leads to a smaller cost for purchasing energy.

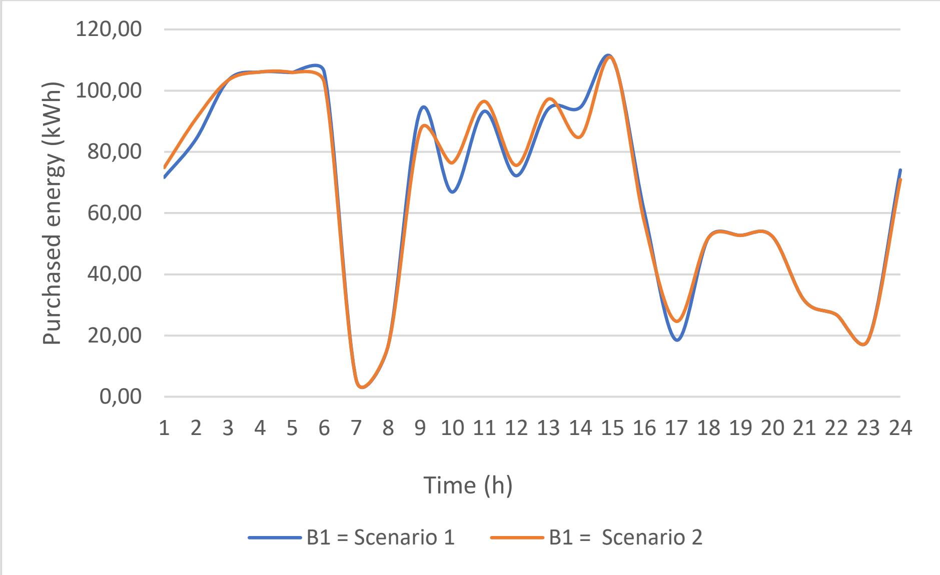

Figure A.3 - Purchased energy by the 30 % of fleet considering DR for both scenarios, in kWh

Table A.2 - Average cost of purchasing energy for both scenarios, in €/MWh

Scenario 1	Scenario 2
30,96	29,75

A.2 Q-learning parameters

This appendix serves to demonstrate the influence of variations in the parameters regarding the Q-learning formulation. The influence of the penalty term is also demonstrated. All tests were performed merely to the 30 % DR penetration level.

A.2.1. α and ε

As to observe the influence of the parameters α and ε throughout the iterative process, two different scenarios were considered.

In the first scenario, α is equal to 0.6 and ε equal to 0.8 during the entire iterative process. In the second scenario, α is equal to 0.6 and its value diminishes to 0.2 throughout the iterative process, while ε diminishes from 0.8 to 0.1 during the iterative process.

Figure A.4 presents the comparison between both scenarios. As it can be seen, diminishing the values of the two variables during the iterative process lead to better results in which concerns to find the optimal results. This happens especially due to the fact that diminishing ε leads the algorithm to exploiting over exploring, and from iteration 300 to 500 agents opt by the most profitable actions most of the times.

A.2.2. Penalty

As to evaluate the effect of the penalty term *w* in defining the optimal solution, Figure A.5 presents the evolution of the reward of the 30 % of the fleet participating in the DR program of this work throughout the iterative process for three different values for *w*. As it can be seen, independently from the value of *w*, the reward converges to approximately the same value, diminishing the total reward as *w* grows due to the fact that deviations between the expected amount of energy to purchase and the actual value of bought energy are more penalized when *w* has a higher value.

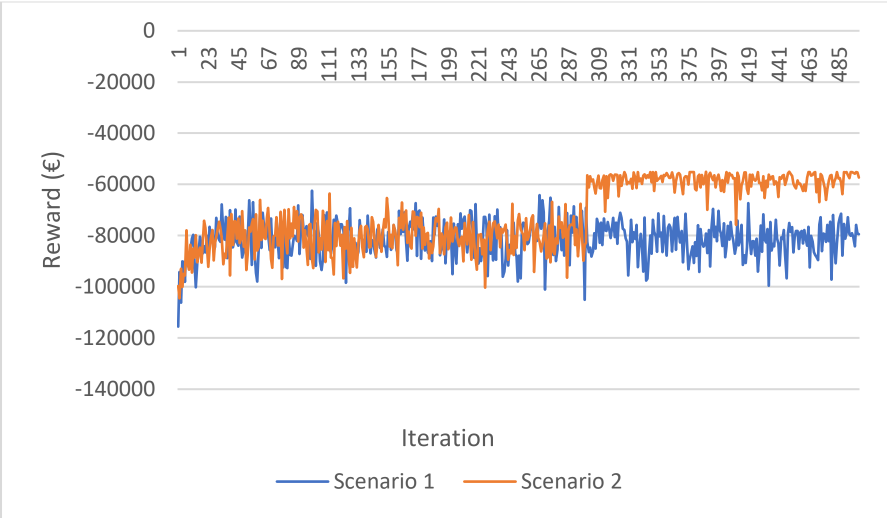

Figure A.4 - Evolution of the reward for both considered scenarios

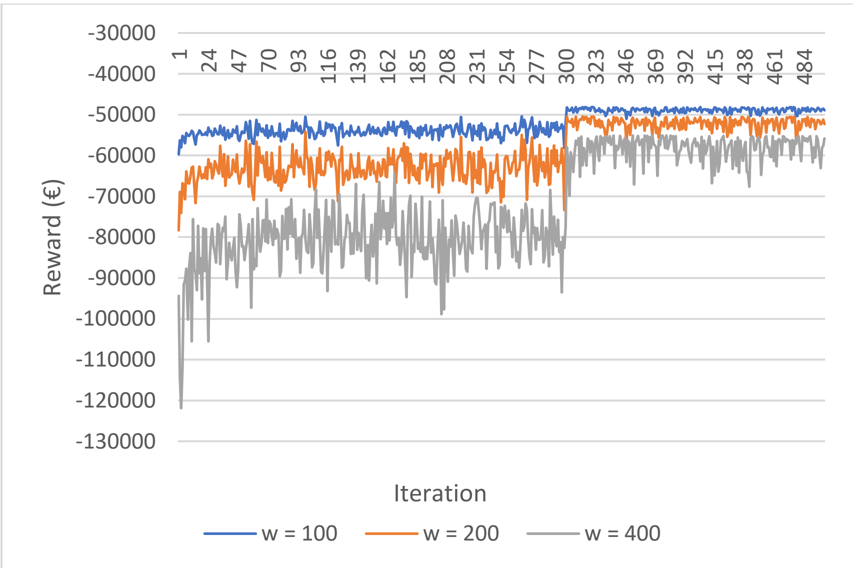

Figure A.5 - Evolution of the reward considering three different values for the penalty term